

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
16 January 2003 (16.01.2003)

PCT

(10) International Publication Number  
**WO 03/004599 A2**

(51) International Patent Classification?: **C12N**  
(21) International Application Number: **PCT/IL02/00540**  
(22) International Filing Date: **2 July 2002 (02.07.2002)**  
(25) Filing Language: **English**  
(26) Publication Language: **English**

(30) Priority Data:  
**60/301,805** **2 July 2001 (02.07.2001)** **US**

(71) Applicant (for all designated States except US): **INSIGHT STRATEGY & MARKETING LTD.** [IL/IL]; Rabin Science Park, P.O. Box 2128, 76 121 Rehovot (IL).

(81) Designated States (national): AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **PELEG, Yoav** [IL/IL]; 18 Neve Alon Street, 76 495 Rehovot (IL). **PANCER, Zeev** [IL/IL]; 3/22 Rav Ashi Street, 69 395 Tel Aviv (IL).

(74) Agent: **G. E. EHRLICH (1995) LTD.**; 28 Bezalel Street, 52 521 Ramat Gan (IL).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: **PROKARYOTIC EXPRESSION CONSTRUCTS, METHODS OF GENERATING SAME AND METHODS OF USING SAME FOR EXPRESSION OF RECOMBINANT PROTEINS IN PROKARYOTIC EXPRESSION SYSTEMS**

(57) Abstract: A method of producing a fusion polypeptide in, and purifying the fusion polypeptide from, bacteria, the method comprises introducing into the bacteria an expression construct encoding the fusion polypeptide which comprises a TAT-derived peptide, and a protein-of-interest, whereby the TAT-derived peptide serves for transport of the fusion polypeptide from the bacterial cytoplasm to the periplasm. The fusion polypeptide is then substantially exclusively purified from the bacterial periplasm.

WO 03/004599 A2

**BEST AVAILABLE COPY**

1  
PROKARYOTIC EXPRESSION CONSTRUCTS, METHODS OF  
GENERATING SAME AND METHODS OF USING SAME FOR  
EXPRESSION OF RECOMBINANT PROTEINS IN PROKARYOTIC  
EXPRESSION SYSTEMS

5

FIELD AND BACKGROUND OF THE INVENTION

The present invention relates to novel prokaryotic expression constructs, methods of generating same and methods of using same for expression of recombinant proteins in prokaryotic expression systems. The prokaryotic expression constructs of the invention induce periplasmic translocation of the expressed proteins and increase recombinant protein expression and bacterial cell mass, resulting in larger quantities of recombinant protein production and substantial simplification of their isolation and purification to homogeneity.

15 Heterologous protein expression has been a technique employed for the past 3 decades, with bacterial expression being recognized as an efficient means of eukaryotic recombinant protein production (Carbon, J., (1993) Genes, Replicators and Centromeres: The First Artificial Chromosomes, in The Early Days of Yeast Genetics. Cold Spring Harbor Laboratory Press pp. 375-390 and Baneyx F., (1999), Recombinant protein expression in *Escherichia coli* Current Opinion in Biotechnology 10: 411-421).

25 These discoveries heralded the creation of a series of prokaryotic expression constructs for use in bacterial systems for recombinant protein production (Studier, F.W. et al. (1990) Use of T7 RNA polymerase to direct expression of cloned genes, Methods in Enzymol. 185, 60-89; Stader, J.A. and Silhavy, T.J. (1990) Engineering *E. coli* to secrete heterologous gene products. Methods in Enzymol. 185, 166-187) with *Escherichia coli* serving as the organism of choice for heterologous protein expression.

*Recombinant protein expression in E. coli*

30 *Escherichia coli* is considered the leading organism for most scientific and commercial applications of recombinant protein expression. Two major

characteristics make *E. coli* ideally suited as an expression system for many proteins: it is easy to genetically manipulate and it grows quickly in inexpensive media. However, in spite of extensive knowledge of the genetics and molecular biology of *E. coli*, not every gene can be expressed efficiently in this organism (Makrides, "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", Microbiological Reviews, 60(3): 512-538, 1996; Swartz, "Advances in *Escherichia coli* Production of Therapeutic Proteins", Curr. Opin. Biotechnol., 12(2): 195-201, 2001).

This may be, in part, due to the structural features of the gene sequence, the stability and translational efficiency of mRNA, degradation of the protein by the host cell proteases and usage of non-favorable codons for the *E. coli* expression system (Makrides, "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", Microbiological Reviews, 60(3): 512-538, 1996; Swartz, "Advances in *Escherichia coli* Production of Therapeutic Proteins", Curr. Opin. Biotechnol., 12(2): 195-201, 2001).

Large-scale protein expression in *E. coli* typically entails bacterial cell growth to high density followed by induction or derepression of the recombinant gene promoter. Tight regulation of the promoter is essential for the synthesis of proteins that may be detrimental to host cells (Makrides, "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", Microbiological Reviews, 60(3):512-538, 1996; Swartz, "Advances in *Escherichia coli* Production of Therapeutic Proteins", Curr. Opin. Biotechnol., 12(2):195-201, 2001). Widely used induction systems that enable better regulation of specific promoters comprise: thermal induction (pL promoter), chemical induction (trp, tac, trc promoters) or nutritional induction (phoA promoter).

Proteins expressed in *E. coli* can be directed either to the cytoplasm or periplasm, or secreted to the extracellular medium (Makrides, "Strategies for Achieving High-Level Expression of Genes in *Escherichia coli*", Microbiological Reviews, 60(3):512-538, 1996, Swartz, "Advances in

*Escherichia coli* Production of Therapeutic Proteins", Curr. Opin. Biotechnol., 12(2):195-201, 2001).

Obtaining soluble, correctly folded proteins is a primary goal in heterologous protein production, yet when *E. coli* is used as an expression system, accumulation of proteins either in the cytoplasm or periplasm of *E. coli*, often occurs inside inclusion bodies, and proteins are often improperly folded.

Protein transport to the bacterial periplasm is a complex and incompletely understood process. The transport of protein through the inner membrane to the periplasmic space requires the inclusion of a signal peptide. The presence of a signal peptide, however, does not always ensure efficient protein translocation through the inner membrane. Periplasmic protein expression and accumulation offers several advantages for heterologous protein production: (i) the oxidizing environment of the periplasm often facilitates the proper folding of proteins; (ii) due to the relatively small number of *E. coli* proteins located in the periplasmic space, purification of expressed proteins is likely to be easier than their purification from the cytoplasm; and (iii) an authentic N terminus can be obtained following *in-vivo* cleavage of the signal peptide during translocation to the periplasm. Thus methods targeting expressed proteins to the periplasmic space will impact proper folding and yields of heterologously expressed proteins and provide a more efficient means of recombinant protein production.

***The Human immunodeficiency virus type 1 transactivating regulatory protein***

The Human Immunodeficiency Virus type 1 transactivating regulatory protein (HIV1 TAT) is an 86-amino acid RNA binding protein involved in replication of the virus. The TAT protein acts in concert with eukaryotic cellular proteins to greatly increase expression of viral genes. The TAT protein increases viral gene expression by acting on the elongation step of transcription. This occurs through the interaction of TAT with a cellular

protein kinase complex known as TAK (also known as cyclin T1/P-TEFb). TAT interacts directly with cyclin T1 and recruits the TAK complex to the viral promoter. TAK hyperphosphorylates a region of RNA polymerase (RNAP) II known as the carboxyl-terminal domain (CTD) which is required for TAT transactivation. Phosphorylation of the CTD is thought to regulate the elongation activity of RNAP II. Therefore hyperphosphorylation of the CTD by TAK is believed to promote the formation of highly processive elongation complexes and provide an explanation for the molecular mechanism of the TAT transactivation function (Yang X, et al (1997) TAK, an HIV-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. Proc. Natl. Acad.Sci. USA 94: 12331-12336; and Herrmann CH and Rice AP. (1995), Lentivirus TAT proteins specifically associate with a cellular protein kinase, TAK, that phosphorylates the carboxyl-terminal domain of the large subunit of RNA polymerase II: a candidate for a TAT cofactor J. Virol. 69: 1612-1620).

The HIV1 TAT gene has two exons. TAT amino acids 1-72 are encoded by exon 1, and amino acids 73-86 are encoded by exon 2. The full-length TAT protein contains a basic region and a cysteine-rich region. The basic region (i.e., amino acids 49-57) is thought to be important for nuclear localization (Ruben, S. et al., (1989) J. Virol. 63: 1-8; Hauber, J. et al., (1989) J. Virol. 63 1181-1187). The cysteine-rich region mediates the formation of metal-linked dimers *in vitro* (Frankel, A. D. et al, (1988) Science 240: 70-73; Frankel, A. D. et al., (1988) Proc. Natl. Acad. Sci USA 85: 6297-6300) and is essential for its activity as a transactivator (Garcia, J. A. et al., (1988) EMBO J. 7: 3143; and Sadaie, M. R. et al., (1989) J. Virol. 63: 1). As in other regulatory proteins, the N-terminal region may be involved in protection against intracellular proteases (Bachmair, A. et al., (1989) Cell 56: 1019-1032).

Purified human immunodeficiency virus type-1 ("HIV") TAT protein is taken up from the surrounding medium by human cells growing in culture (A. D. Frankel and C. O. Pabo, (1988) "Cellular Uptake of the TAT Protein from Human Immunodeficiency Virus", *Cell*, 55, pp. 1189-93).

- 5 A peptide derived from the basic region of the TAT protein has been shown to function as a transduction domain for the delivery of biologically active molecules into the cytoplasm and nuclei of eukaryotic cells, *in vitro* and *in vivo*. Use of synthetic or recombinant peptides consisting of all or part of the amino acid sequence YGRKKRRQRRR (amino acids 47 to 57 in the HIV1
- 10 TAT protein), covalently attached to cargo molecules such as polypeptides and nucleic acids (Fawell et al, , (1994) "TAT-mediated Delivery of Heterologous Proteins into Cells", *PNAS*, 91(2): 664-668; Vives et al, (1997) "A truncated HIV-1 TAT Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus", *J. Biol. Chem.*, 272(25):
- 15 16010-16017,; Torchilin et al, (2001) "TAT Peptide on the Surface of Liposomes Affords their Intracellular Delivery Even at Low Temperature and in the Presence of Metabolic Inhibitors", *PNAS*, 98(15): 8786-8791) can transport these molecules into the cytoplasm and nuclei of eukaryotic cells. Minimal requirements for eukaryotic transport comprise the TAT basic region,
- 20 or derivatives thereof, such as, for example, a TAT peptide flanked by two Glycine residues (Fawell et al, (1994) "TAT-mediated Delivery of Heterologous Proteins into Cells", *PNAS*, 91(2): 664-668,; Vives et al, (1997) "A truncated HIV-1 TAT Protein Basic Domain Rapidly Translocates through the Plasma Membrane and Accumulates in the Cell Nucleus", *J. Biol. Chem.*,
- 25 272(25):16010-16017,; Torchilin et al, (2001) "TAT Peptide on the Surface of Liposomes Affords their Intracellular Delivery Even at Low Temperature and in the Presence of Metabolic Inhibitors", *PNAS*, 98(15):8786-8791, and U.S. Pat. No. 5,652,122; U.S. Pat. No. 5,670,617; U.S. Pat. No. 5,674,980; U.S. Pat. No. 5,747,641 U.S. Pat. No. 5,804,604; U.S. Pat. No. 6,221,355).

These references disclose TAT-mediated transport of proteins in eukaryotic systems. The references disclose heterologous protein production covalently attached to a segment of the TAT protein in *E. coli*, and subsequent delivery of this construct as a means of delivering TAT-conjugated proteins into eukaryotic cell cytoplasm or nuclei. Bacterial subcellular localization was not addressed, nor was there an attempt to address TAT-mediated effects on increased bacterial protein production. Finally, the ability of the TAT-containing constructs to undergo cleavage of the N-terminal TAT segment, resulting in properly folded, mature heterologously expressed protein is a unique application and novel development of the present invention alone. All the above references result in TAT-conjugated products being isolated from bacterial expression systems and used for delivery in eukaryotic systems. Therefore, to date, no reference has been made to any effect of TAT peptides or derivatives in protein transport and processing in prokaryotic cells.

While there are clear advantages to the use of current recombinant protein production techniques in *E. coli*, many heterologous polypeptides fail to fold into their native state when expressed and instead are either degraded by the cellular proteolytic machinery or accumulate in insoluble form, typically as inclusion bodies, and hence applications addressing this issue are of paramount importance. Misfolding is a particularly vexing problem in the expression of mammalian proteins, especially proteins that are composed of multiple subunits, have several disulfide bonds, or containing prosthetic groups. (George Georgiou and Pascal Valax (1996) Expression of correctly folded proteins in *Escherichia coli*, Current Opinion in Biotechnology, 7: 190-197). Misfolding may, in certain circumstances, be overcome, however, it is clear that existing methodologies do not allow for ease of expression of properly folded heterologous proteins, or of many eukaryotic proteins at all.

Thus there is a widely recognized need for, and it would be highly advantageous to have, a prokaryotic expression system devoid of the above limitations. Expression systems yielding high levels of properly folded and

processed, functional recombinant proteins enabling purification with relative ease, in a cost-efficient setting is highly desirable, providing a clear and direct impact on the treatment of an infinite number of pathologies and diseases, and serving a multitude of other applications.

5

### SUMMARY OF THE INVENTION

While reducing the present invention to practice it was unexpectedly found that incorporation of a TAT-derived peptide in a fusion polypeptide containing a protein-of-interest is sufficient to enable periplasmic translocation of the fusion polypeptide, where efficient isolation of the fusion polypeptide was readily accomplished.

More surprisingly, incorporation of a bacterial signal sequence in frame with the TAT-derived peptide and protein-of-interest provided for cleavage of the signal sequence and the TAT-derived sequence, and provided a mature, properly folded and functional protein, readily isolatable and purifiable. The signal sequence thus functioning as providing a bacterial protease cleavage recognition sequence for cleavage of the periplasmic targeting, or TAT-derived peptide, and the cleavage sequence itself. Once cleaved, the remaining protein-of-interest folds properly within the oxidizing environment provided by the bacterial periplasm.

Hence, the present invention relates to novel prokaryotic expression constructs, methods of generating same and methods of using same for expression of recombinant proteins in prokaryotic expression systems. The prokaryotic expression constructs of the present invention encode for a TAT-derived polypeptide fused to a protein-of-interest, targeting the resulting recombinant fusion polypeptide to the periplasm. Incorporation of a bacterial signal sequence, or other bacterial protease cleavage recognition sequence, within the construct facilitates TAT-derived polypeptide cleavage, releasing the mature recombinant protein-of-interest. Within the periplasm, the mature recombinant protein-of-interest can undergo proper folding, and can be



isolated therefrom to homogeneity with relative ease, and minimum expense. Non-bacterial protease cleavage recognition sequence can be used for post isolation processing of the fusion polypeptide into the mature recombinant protein-of-interest.

5       The prokaryotic expression constructs of the invention increase recombinant protein expression, periplasmic translocation of the expressed proteins and also increases bacterial cell mass.

Thus, the present invention provides processes and products for the efficient production of mature recombinant proteins-of-interest using  
10       prokaryotic expression systems. The method and products provided herein are novel and/or beneficial in that they provide for (i) high levels of expression of a protein-of-interest in prokaryotic cells; (ii) proper folding of a protein-of interest once expressed; and/or (iii) ease of isolation of the proteins-of-interest from a bacterial periplasm.

15       It is one object of the present invention to provide methods, expression constructs and kits for producing a protein-of-interest in, and purifying the protein-of-interest substantially, exclusively from a bacterial periplasm.

It is another object of the present invention to provide methods, expression constructs and kits for producing a fusion polypeptide in, and  
20       purifying the fusion polypeptide substantially, exclusively from a bacterial periplasm.

It is still another object of the present invention to provide assays for determining whether inclusion of a certain TAT-derived sequence within a nucleic acid expression construct is sufficient for effective periplasmic  
25       targeting.

According to one aspect of the present invention there is therefore provided an assay of determining whether a TAT-derived peptide is an effective periplasmic targeting sequence. The assay comprises introducing into bacteria an expression construct encoding a fusion polypeptide comprising the  
30       TAT-derived peptide and a reporter protein, and determining to what extent the

fusion polypeptide accumulates within the periplasm, thereby determining whether the TAT-derived peptide is an effective periplasmic targeting sequence. Preferably, the expression construct further harbors a reporter sequence for ease of identification of expression within a bacterial periplasm.

5 This assay can be used, by one ordinarily skilled in the art, to determine which of the homologous TAT-derived sequences from, for example, any of the viral species listed below, and/or modifications thereof is most efficient for the purpose of implementing the various methods, constructs and kits of the present invention which are further described hereinafter.

10 According to another aspect of the present invention there is provided a method of producing a protein-of-interest in, and purifying the protein-of-interest from, bacteria. The method comprises introducing into the bacteria an expression construct encoding a fusion polypeptide, which comprises a TAT-derived peptide, a signal sequence and the protein-of-interest. The TAT-  
15 derived peptide serves for transport of the fusion polypeptide from the bacterial cytoplasm to the periplasm, and the signal sequence facilitates processing the fusion polypeptide to a mature protein, consisting essentially of the protein-of-interest and substantially lacking the TAT-derived peptide and signal sequence. The mature protein is then substantially exclusively purified  
20 from the bacterial periplasm.

According to another aspect of the present invention there is provided a method of producing a fusion polypeptide in, and purifying the fusion polypeptide from, bacteria. The method comprises introducing into the bacteria an expression construct encoding the fusion polypeptide, which  
25 comprises a TAT-derived peptide, and a protein-of-interest. The TAT-derived peptide serves for transport of the fusion polypeptide from the bacterial cytoplasm to the periplasm. The fusion polypeptide is then substantially exclusively purified from the bacterial periplasm.

In one embodiment of the present invention, the fusion polypeptide  
30 further comprises a protease cleavage recognition sequence positioned

between the TAT-derived peptide and the protein-of-interest, and the method further comprises cleaving the fusion polypeptide with a protease specific to the protease cleavage recognition sequence, releasing the protein-of interest pre or post purification.

5 In another embodiment of the present invention, purifying the fusion polypeptide or the protein-of interest substantially exclusively from the bacterial periplasm comprises (a) isolating the bacterial periplasmic compartment from other subcellular compartments; (b) lysing the bacterial periplasmic compartment; and (c) purifying the protein of interest or fusion  
10 polypeptide. Preferably, isolating the bacterial periplasmic compartment utilizes the methods of subcellular fractionation, differential gradient centrifugation and/or gel electrophoresis. Preferably, purifying the fusion polypeptide or protein-of-interest utilizes methods of purification and analysis such as, but not limited to, column chromatography, electrophoresis, filtration,  
15 ultrafiltration, gradient centrifugation, HPLC, Western blot analysis, mass spectroscopy, GLC, and/or immunocytochemistry.

According to still another aspect of the present invention there is provided a prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a signal sequence and a protein-of-interest,  
20 wherein the TAT-derived peptide serves for transport of the fusion polypeptide to the periplasm of the prokaryotic cell and the signal sequence facilitates processing of the fusion polypeptide to yield a mature protein consisting essentially of the protein-of-interest and lacking the TAT-derived peptide and said signal sequence.

25 According to still another aspect of the present invention there is provided a prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a protease cleavage recognition sequence and a protein-of-interest, wherein the protease cleavage recognition sequence is positioned between the TAT-derived peptide and the protein-of-interest,

whereby the TAT-derived peptide serves for transport of the fusion polypeptide to the periplasm of the prokaryotic cell.

According to still another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide  
5 encoding a TAT-derived peptide, a second polynucleotide harboring an intact polylinker cloning sequence being operably linked to the first polynucleotide and a third polynucleotide harboring a prokaryotic promoter, being operably linked to the first polynucleotide.

According to a described preferred embodiment the nucleic acid  
10 expression construct further comprises a polynucleotide encoding a protease cleavage recognition sequence in frame with the TAT-derived peptide.

According to still another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a signal  
15 sequence in frame with the TAT-derived peptide, and a third polynucleotide harboring a polylinker cloning sequence, being operably linked to the second polynucleotide and a fourth polynucleotide harboring a prokaryotic promoter, being operably linked to the first polynucleotide.

According to still another aspect of the present invention there is  
20 provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a protease cleavage recognition sequence in frame with the TAT-derived peptide, a third polynucleotide encoding a protein-of-interest in frame with the protease cleavage recognition sequence and a fourth polynucleotide harboring a  
25 prokaryotic promoter, being operably linked to the first polynucleotide.

According to still another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a signal  
30 sequence in frame with the TAT-derived peptide, a third polynucleotide encoding a protein-of-interest in frame with the signal sequence and a fourth

polynucleotide harboring a prokaryotic promoter, being operably linked to the first polynucleotide.

According to still another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a mammalian secreted protein-of-interest in frame with the TAT-derived peptide and a third polynucleotide harboring a reporter gene, being operably linked to the second polynucleotide.

According to still another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a mammalian, non-nuclear, protein-of-interest in frame with the TAT-derived peptide, and a third polynucleotide harboring a reporter gene, being operably linked to the second polynucleotide.

Further according to the present invention there are provided kits comprising any of the above expression constructs and optionally reagents required for bacterial transformation and or transfection, including, for example, buffers, competent bacteria, phage packaging proteins, helper phages, enzymes, such as restriction endonucleases, ligases and DNA polymerases, and the like.

According to a still further aspect of the present invention there is provided a prokaryotic cell engineered to express the expression constructs described herein.

According to features of the described preferred embodiments of the present invention the prokaryotic cell utilized for expression of the fusion polypeptide, protein-of-interest and/or bacterial expression construct is a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*,

*Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

According to further features of the described preferred embodiments of the present invention the TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18), or alternatively, is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

According to still further features of the described preferred embodiments of the present invention the TAT-derived peptide is N-terminal to the protein-of-interest in fusion polypeptides obtained from any of the above listed expression constructs encoding fusion polypeptides comprising a TAT-derived peptide and a protein-of-interest.

According to still further features of the described preferred embodiments of the present invention expression constructs containing polynucleotide sequences encoding the TAT-derived peptide are located upstream to sequences encoding the protein-of-interest.

According to still further features of the described preferred embodiments of the present invention the TAT-derived peptide is N-terminal to the signal sequence in fusion polypeptides obtained from any of the above listed expression constructs encoding fusion polypeptides comprising a TAT-derived peptide and a signal sequence.

According to still further features of the described preferred embodiments of the present invention expression constructs containing polynucleotide sequences encoding the TAT-derived peptide are located upstream to sequences encoding the signal sequence.

According to still further features of the described preferred embodiments of the present invention the signal sequence comprises a

positively charged amino-terminus, a hydrophobic central region, and a neutral but polar carboxy-terminus.

According to still further features of the described preferred embodiments of the present invention, the promoter sequence operably linked  
5 to the polynucleotides encoding the fusion polypeptide are either constitutive or inducible, and provide for low or high level expression of the construct.

According to still further features of the described preferred embodiments of the present invention, the expression constructs contain a reporter gene, such as  $\beta$ -galactosidase, chloramphenicol acetyl transferase,  
10 luciferase and a fluorescent protein.

According to still other features of the described preferred embodiments of the present invention, the expression constructs further comprise a polynucleotide encoding a positive or a negative selection marker.

According to still further features of the described preferred  
15 embodiments of the present invention the protein-of-interest may be selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an  
20 amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic  
25 erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a  
30 streptavidin.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described below. In case of conflict, the patent specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

#### 10 BRIEF DESCRIPTION OF THE DRAWINGS

The invention is herein described, by way of example only, with reference to the accompanying figures. With specific reference now to the figures in detail, it is stressed that the particulars shown are by way of example and for purposes of illustrative discussion of the preferred embodiments of the present invention only, and are presented in the cause of providing what is believed to be the most useful and readily understood description of the principles and conceptual aspects of the invention. In this regard, no attempt is made to show structural details of the invention in more detail than is necessary for a fundamental understanding of the invention, the description taken with the drawings making apparent to those skilled in the art how the several forms of the invention may be embodied in practice.

In the drawings:

FIG. 1 is a schematic representation of the expression construct pTAT-STB-hGH. Relative positions of some functional and regulatory elements are shown in the diagram, including: Kan - Kanamycin-resistance gene; lacI - lacI repressor; ori - origin of replication; TAT-STB-hGH - TAT-derived sequence fused to *E. coli* heat stable enterotoxin II signal peptide and to human growth hormone.

FIG. 2A is a schematic representation of the expression cassette. KpnI-SacI sites flank the expression cassette. Relative positions of some restriction



sites, functional and regulatory elements are shown. Arrows indicate the direction of transcription from the heat-inducible pL promoter and from the constitutive promoter. Positioning of the multiple cloning site (MCS); and transcriptional termination (TT) site are as indicated.

5        FIG. 2B is a schematic representation of the expression construct hGHTSP/pACYC184. Relative positions of some functional and regulatory elements are as indicated, including: Tc - Tetracycline-resistance gene; ori-origin of replication; repressor - thermo-labile repressor; TSP-hGH - TAT-derived sequence fused to *E. coli* heat stable enterotoxin II signal peptide and  
10      to human growth hormone.

FIG. 3 is a photograph of a protein blot stained with coomassie blue revealing cytoplasmic accumulation of TAT-STB-hGH, and periplasmic accumulation of mature hGH, expressed in bacteria harboring constructs with the TAT-derived peptide. Large asterisks mark the non-processed fusion hGH  
15      and small asterisks mark the mature hGH. Three different *E. coli* isolates harboring the pTAT-STB-hGH plasmid were examined, with and without induction, and sizes were compared to the protein molecular weight marker (Prestained Protein Marker, Broad Range, Cat. No.: P7708S, New England BioLabs), and presented in kilo-daltons (kDa).

20        FIG. 4 is a photograph of a protein blot probed with anti-hGH antibodies for the detection of recombinant hGH fusion polypeptides. The upper panel shows positive staining for cytoplasmic and periplasmic proteins of 22-25 kDa in size. The lower panel shows positive staining for proteins of 22-25 kDa in size, in samples of periplasmic proteins isolated before (-) and  
25      after (+) induction. Size differences reflect the periplasmic accumulation of mature hGH, as well as non-processed TAT-hGH, in respective samples.

FIG. 5 is a graph illustrating the growth rates of *E. coli* strains transformed with constructs with and without the TAT-derived sequence. *E. coli* MM294 harboring expression constructs hGHSP/pACYC184 and  
30      hGHTSP/pACYC184, (without and with the TAT-derived sequence,

respectively) were grown in a 5 liter fermenter. Samples were taken as indicated, for measurements of optical density (OD) at 600 nm. Open and closed circles represent growth of *E. coli* harboring expression construct hGHTSP/pACYC184 and hGHSP/pACYC184, respectively. Arrow indicates  
5 time of induction, when the culture reached 42 °C.

### DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is of (i) novel prokaryotic expression constructs; (ii) methods of generating the expression constructs; and (iii) methods of using  
10 the expression constructs for the expression of recombinant proteins-of-interest in prokaryotic expression systems. Specifically, the prokaryotic expression constructs encode for fusion polypeptides which comprise a TAT-derived peptide, and optionally a prokaryotic signal sequence, or a protease cleavage recognition sequence, fused to a molecule of interest, resulting in the targeting  
15 of the recombinant fusion polypeptide to the bacterial periplasm, wherein optionally the TAT-derived peptide is cleaved and the remaining protein-of-interest undergoes proper folding, resulting in the production of a mature recombinant protein-of-interest.

The principles and operation of the present invention may be better  
20 understood with reference to the drawings and accompanying descriptions.

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable  
25 of other embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for the purpose of description and should not be regarded as limiting.

The characterization of a novel periplasmic transport mechanism in prokaryotes for heterologously expressed, mature proteins-of-interest is  
30 described herein. The transport mechanism utilizes the incorporation of a

genetic fusion of a TAT-derived peptide and a protein-of-interest, to target the translated fusion polypeptide to the bacterial periplasm. Incorporation of a bacterial signal sequence, functioning as a bacterial protease cleavage recognition sequence, facilitates cleavage of the signal/cleavage sequence, as well as cleavage of the TAT-derived peptide. Targeted proteins undergo cleavage in the bacterial periplasm, liberating the protein-of-interest, enabling its proper folding and ultimate isolation as a mature protein-of-interest, devoid of the TAT-derived peptide and bacterial signal sequence from the bacterial periplasmic space.

The TAT-derived peptide utilized in these studies comprises amino acids 47 to 57 of the TAT protein of HIV-1 (YGRKKRRQRRR) (SEQ ID 18). It is to be understood, and as is further defined hereinunder, that other sequences of any TAT protein may be similarly utilized to target fusion polypeptides containing a protein-of-interest to a bacterial periplasm. Also, man or naturally modified sequences can be employed, such modified sequences may include, for example, amino acids similar to those amino acids naturally present in the TAT protein (e.g., basic and hydrophylic amino acids), the absence of at least one amino acid and the addition of at least one amino acid. TAT proteins have been identified in several viral species including, but not limited to, human immunodeficiency virus type 1 (HIV-1), human immunodeficiency virus type 2 (HIV-2), equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus. Certain sequence modifications to TAT sequences, which can be used in context of the present invention are described in Wender, PA (2000) The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. PNAS. 97(24)13003-13008; and Futaki S (2001) Arginine-rich peptides. JBC 276(8)5836-5840.).

Amino acid substitutions are typically of single residues; insertions usually will be on the order of from about 1 to 5 amino acids, although

considerably larger insertions may be tolerated. Deletions range from about 1 to about 2 residues, although in some cases deletions may be larger.

Substitutions, deletions, insertions or any combination thereof may be used to arrive at a final derivative. Generally these changes are done on a few amino acids to minimize the alteration of the molecule. However, larger changes may be tolerated in certain circumstances. When small alterations in the characteristics of TAT derived peptide are desired, substitutions are generally made in accordance with the following chart:

10	Original Residue	Exemplary Substitutions
	Ala	Ser
	Arg	Lys
	Asn	Gln, His
	Asp	Glu
15	Cys	Ser
	Gln	Asn
	Glu	Asp
	Gly	Pro
	His	Asn, Gln
20	Ile	Leu, Val
	Leu	Ile, Val
	Lys	Arg, Gln, Glu
	Met	Leu, Ile
	Phe	Met, Leu, Tyr
25	Ser	Thr
	Thr	Ser
	Trp	Tyr
	Tyr	Trp, Phe
	Val	Ile, Leu

In one aspect of the present invention there is disclosed an assay whereby TAT-derived peptides may be evaluated for their potential as periplasmic targeting sequences, and will be further discussed hereinbelow.

Incorporation of the TAT-derived peptide was sufficient to direct  
5 heterologously expressed fusion polypeptides to the bacterial periplasm. Western blots of human growth hormone (hGH) - TAT fusion polypeptides readily localized to bacterial periplasms. Incorporation of a signal peptide, in this case the heat-stable enterotoxin II signal peptide (STB), further resulted in processing of the fusion polypeptide to yield the mature, appropriately sized  
10 (22 kDa) hormone, as opposed to the fusion peptide isolated by the TAT-hGH construct alone (24 kDa). Both constructs provided greater amounts of hGH as compared to constructs harboring STB-hGH alone, without inclusion of the TAT-derived peptide.

Incorporation of the TAT-derived peptide sequence upstream of the  
15 signal peptide resulted in production of the properly processed, mature hGH protein, as determined by N-terminal protein sequencing of isolated proteins, which was identical to the published sequence of the mature hGH protein.

Moreover, higher levels of growth were sustained for a significantly longer period of time in bacterial strains harboring expression constructs  
20 containing the TAT-derived peptide and hGH, as compared to constructs harboring hGH alone.

Growth differences were observed primarily, and most unexpectedly during the induction phase, and not the growth phase. ELISA analysis confirmed data from the growth curves, indicating almost twice the amount of  
25 hGH protein produced in strains harboring the TAT-derived peptide.

The study presented herein is therefore the first to demonstrate that a viral derived protein product, TAT, directs protein translocation to the periplasmic space when the TAT-derived peptide is produced in frame with the protein-of-interest.

Furthermore, the study presented herein is the first to describe the incorporation of bacterial signal sequence, in this case functioning as a bacterial protease recognition sequence, that facilitates cleavage of the TAT-derived peptide and signal sequence, releasing a mature recombinant protein-of-interest.

Translocation of the mature protein-of-interest within the periplasmic space facilitated proper protein folding and enabled protein function, and thus the present invention provides a novel and highly efficient means of high yield heterologous expression and ease of purification of a protein-of-interest in a bacterial cell.

As used herein the phrase "mature protein-of-interest" includes heterologously expressed proteins that are processed and folded to assume a final sequence and three-dimensional structure equal or similar to published accounts of the protein.

As used herein the phrase "expression construct" includes nucleic acid vectors that contain gene sequences essential for maintenance and propagation of the vector and directing/regulating the transcription/translation of inserted or subcloned sequences of interest.

As used herein the phrases "fusion polypeptide" and "fusion protein" may be used interchangeably and refer to a genetically engineered covalently linked protein derived from the joint, in-frame expression, of two or more heterologous nucleic acid sequences.

As used herein in the specification and in the claims section that follows, the phrase "signal sequence" refers to a short (e.g., 15-40) amino acid sequences, which allow proteins to transport through the bacterial inner membrane to the periplasm. During transport of proteins out of the cytoplasm, the signal peptide is typically removed by signal peptidases, thereby releasing a mature protein at the desired non-cytoplasmic location.

As used herein the term "periplasm" is understood to be the space between the bacterial inner and outer membranes.

As used herein the phrase "TAT-derived peptide" encompasses naturally appearing and man modified peptide sequences derived from any TAT protein of any virus whose genome encodes for a TAT protein and that is positively functional in the assay of determining whether a TAT-derived peptide is an effective periplasmic targeting sequence. Modifications to a naturally occurring TAT-derived peptide may include addition, deletion or substitution of one or more amino acids, as well known in the art. The periplasmic targeting activity of each such modified sequence can be determined via the assay described herein.

Thus, according to one aspect of the present invention there is provided an assay of determining whether a TAT-derived peptide is an effective periplasmic targeting sequence. The assay according to this aspect of the invention comprises introducing into bacteria an expression construct encoding a fusion polypeptide comprising the TAT-derived peptide and a reporter protein and determining to what extent said fusion polypeptide accumulates within the periplasm, thereby determining whether the TAT-derived peptide is an effective periplasmic targeting sequence. Experimental methods effective in implementing the above assay are described in the preferred embodiments and Examples sections of this application and/or are otherwise well known to the skilled artisan. The assay according to this aspect of the present invention is useful in identifying novel, naturally occurring and/or man modified TAT-derived sequences, effective in periplasmic targeting of proteins fused thereto and hence useful in context of the various embodiments and aspects of the present invention.

According to another aspect of the present invention there is provided a method of producing a protein-of-interest in, and purifying the protein-of-interest from, bacteria. The method comprises introducing an expression construct encoding a fusion polypeptide into the bacteria. The expression construct comprises a TAT-derived peptide, a signal sequence and the protein-of-interest, wherein the TAT-derived peptide serves for transporting the fusion

polypeptide from the cytoplasm to the periplasm of the bacteria and the signal sequence facilitates processing of the fusion polypeptide to a mature protein. The mature protein thus consists essentially of the protein-of-interest and substantially lacks the TAT-derived peptide and signal sequence. The mature protein is then purified substantially exclusively from the bacterial periplasm.

According to another aspect of the present invention there is provided an additional method of producing a protein-of-interest in, and purifying the protein-of-interest from, bacteria. This method comprises introducing an expression construct encoding a fusion polypeptide into the bacteria. The expression construct comprises a TAT-derived peptide, a protease cleavage recognition sequence and the protein-of-interest, wherein the TAT-derived peptide serves for transporting the fusion polypeptide from the cytoplasm to the periplasm of the bacteria. The protease cleavage recognition sequence facilitates processing of the fusion polypeptide to a mature protein-of-interest by either by cleavage of the TAT-derived peptide and the protease cleavage recognition sequence by bacterial proteases prior to isolation from the periplasm, or by cleavage of the TAT-derived peptide and the protease cleavage recognition sequence by a protease post isolation from the periplasm. In any case, the mature protein consists essentially of the protein-of-interest and substantially lacks the TAT-derived peptide and the protease cleavage recognition sequence.

According to another aspect of the present invention there is provided a method of producing a fusion polypeptide in, and purifying a fusion polypeptide from, bacteria. The method comprises introducing into the bacteria an expression construct encoding the fusion polypeptide comprising a TAT-derived peptide, and a protein-of-interest. The TAT-derived peptide serves for transport of the fusion polypeptide from the cytoplasm to the periplasm of the bacteria. Once translocated to the bacterial periplasm, the fusion polypeptide is purified substantially exclusively from the bacterial periplasm.



Inclusion of a TAT-derived peptide in the methods disclosed herein of construction and isolation of fusion polypeptides containing a protein-of-interest, and fusion polypeptides containing a protein-of-interest and signal sequence or a protease cleavage sequences, processed and ultimately lacking the signal sequence and the TAT-derived peptide, facilitates increased yield of  
5 the protein-of-interest.

These methods additionally provide means with which to purify the mature protein-of-interest or fusion polypeptide from the periplasm of prokaryotic cells with higher yields than previously disclosed methodologies.

10 Purification of fusion polypeptides and proteins-of-interest from the bacterial periplasm may be accomplished according to methods well known in the art. For example, purification of fusion polypeptides and proteins-of-interest from the bacterial periplasm may be accomplished by isolating the bacterial periplasmic compartment from other subcellular compartments,  
15 lysing the bacterial periplasmic compartment and purifying the protein of interest from the lysate, using, for example, at least one of the following purification and/or analysis methods: column chromatography, electrophoresis, filtration, ultrafiltration, gradient centrifugation, preparative HPLC, analytic HPLC, Western blot analysis, mass spectroscopy, GLC, and/or  
20 immunocytochemistry.

In order to generate the nucleic acid constructs of the present invention disclosed hereinbelow, polynucleotide segments encoding a TAT-derived peptide, the protein-of-interest and optionally a signal sequence or a protease cleavage recognition sequence can be ligated into commercially available  
25 expression construct systems suitable for transforming bacterial cells and for directing the expression of the fusion polypeptide within the transformed cells. It will be appreciated that such commercially available vector systems can easily be modified via commonly used recombinant techniques in order to replace, duplicate or mutate existing promoter or enhancer sequences and/or  
30 introduce any additional polynucleotide sequences such as for example,

sequences encoding additional selection markers or sequences encoding reporter polypeptides, and as such, encompass preferred embodiments of the present invention.

Suitable bacterial expression constructs for use with the present invention include, but are not limited to the pCAL, pUC, pET, pETBlue™ (Novagen), pBAD, pLEX, pTrcHis2, pSE280, pSE380, pSE420 (Invitrogen), pKK223-2 (Clontech), pTrc99A, pKK223-3, pRIT2T, pMC1871, pEZZ 18 (Pharmacia), pBluescript II SK (Stratagene), pALTER-Ex1, pALTER-Ex2, pGEMEX (Promega), pFivE (MBI), pQE (Qiagen) commercially available expression constructs, and their derivatives. In preferred embodiments of the present invention the construct may also include, a plasmid, a bacmid, a phagemid, a cosmid, or a bacteriophage. Bi-functional or shuttle vectors suitable for propagation and gene expression both in prokaryote and eukaryote organisms are also within the scope of the present invention.

Nucleotide sequences are typically operably linked to, i.e., positioned, to ensure the functioning of an expression control sequence. These expression constructs are typically replicable in the cells either as episomes or as an integral part of the cell's chromosomal DNA, and may contain appropriate origins of replication for the respective prokaryotic strain employed for expression. Commonly, expression constructs contain selection markers, such as for example, tetracycline resistance, ampicillin resistance, kanamycin resistance or chlormaphenicol resistance, facilitating detection and/or selection of those bacterial cells transformed with the desired nucleic acid sequences (see, e.g., U.S. Pat. No. 4,704,362). These markers, however, are not exclusionary, and numerous others may be employed, as known to those skilled in the art. Indeed, in a preferred embodiment of the present invention expression constructs contain both positive and negative selection markers.

Similarly, reporter genes may be incorporated within expression constructs to facilitate identification of transcribed products. Accordingly, in a preferred embodiment of the present invention, reporter genes utilized are

selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein, e.g., green fluorescent protein (GFP).

Prokaryotic promoter sequences regulate expression of the encoded  
5 polynucleotide sequences, and in preferred embodiments of the present invention, are operably linked to polynucleotides encoding the TAT-derived peptide and polynucleotides encoding the protein-of-interest. In additional preferred embodiments of the present invention, these promoters are either constitutive or inducible, and provide a means of high and low levels of  
10 expression of the fusion polypeptides.

Many well-known bacterial promoters, including the T7 promoter system, the lactose promoter system, tryptophan (Trp) promoter system, Trc/Tac promoter systems, beta-lactamase promoter system, tetA promoter systems, arabinose regulated promoter system, Phage T5 promoter, or a  
15 promoter system from phage lambda, may be employed, and others, as well, all comprise preferred embodiments of the present invention. The promoters will typically control expression, optionally with an operator sequence and may include ribosome binding site sequences for example, for initiating and completing transcription and translation.

20 According to additional preferred embodiments, the vector may also contain expression control sequences, enhancers that may regulate the transcriptional activity of the promoter, appropriate restriction sites to facilitate cloning of inserts adjacent to the promoter and other necessary information processing sites, such as RNA splice sites, polyadenylation sites and  
25 transcription termination sequences as well as any other sequence which may facilitate the expression of the inserted nucleic acid.

For proteins normally expressed within a bacterial cell, protein export from the cytoplasm to the periplasmic space as premature forms naturally occurs. These premature proteins have short (15-30) specific amino acid  
30 sequences, which allow proteins to transport through the inner membrane to

the periplasm. These short amino acid sequences are commonly referred to as signal peptides. The premature protein interacts with the bacterial secretion apparatus, and during transport of the proteins out of the cytoplasm, the signal peptide is typically removed by signal peptidase thereby leaving a mature protein at the desired non-cytoplasmic location. The features of a typical bacterial signal peptide include a positively charged amino-terminus (n-region), a hydrophobic central region (h-region), and a neutral but polar carboxy-terminus (c- region). A helix-breaking residue (Pro or Gly) usually marks the boundary between the h-region and the c-region at the -6 (P6) position relative to the cleavage site. The cleavage recognition sequence consists of small residues at the -1 (P1) and -3 (P3) positions relative to the cleavage site (Edman, M., Jarhede, T., Sjostrom, M., and Wieslander, A. (1999) Different sequence patterns in signal peptides in signal peptidases from mycoplasmas, other gram- positive bacteria, and *Escherichia coli*: multivariate data analysis. *Proteins Struc. Func. Gen.* 35: 195-205; Jones, J. D., and Giersch, L. M. (1994) Effect of charged residue substitutions on the membrane-interactive properties of signal sequences of the *Escherichia coli* LamB protein. *Biophys. J.* 67: 1534-1544; and von Heijne, G. (1990) The signal peptide. *J. Membr. Biol.* 115: 195-201).

Only the incorporation of the TAT-derived peptide, as disclosed herein, however, facilitated significant heterologous protein-of-interest transport to the periplasm, and incorporation of a bacterial signal sequence, functioning as a bacterial protease recognition sequence, provided for efficient processing to yield significantly greater quantities of the mature protein-of-interest.

Therefore, according to another aspect of this invention there is provided a prokaryotic nucleic acid expression construct comprising a first polynucleotide encoding a fusion polypeptide comprising a signal sequence, TAT-derived peptide and a protein-of-interest. The TAT-derived peptide serves for transport of the fusion polypeptide to the periplasm of the bacterial cell expressing the construct, and expression of the signal sequence facilitates

cleavage of both the TAT-derived peptide and signal sequence, resulting in a mature protein-of-interest, devoid of the TAT-derived peptide and signal sequence. The construct also comprises a second polynucleotide harboring a prokaryotic promoter, being operably linked to the first polynucleotide, providing a means for regulating expression of the construct. In preferred  
5 embodiments of the invention, the construct additionally comprises a reporter gene and positive and/or negative selection markers.

Since the signal sequence functions to cleave both the signal sequence itself, and the TAT-derived peptide from the protein-of-interest, it is apparent  
10 that the signal sequence is functioning as a bacterial protease cleavage recognition sequence.

Accordingly, in another aspect of the present invention there is provided a nucleic acid expression construct comprising a polynucleotide encoding a TAT-derived peptide, a polynucleotide encoding a protease cleavage  
15 recognition sequence in frame with the TAT-derived peptide, a polynucleotide encoding a protein-of-interest in frame with the protease cleavage recognition sequence and a polynucleotide harboring a prokaryotic promoter, being operably linked to the polynucleotide encoding the TAT-derived peptide, providing a means for regulating expression of the construct. In preferred  
20 embodiments of the invention, the construct similarly additionally comprises a reporter gene and positive and/or negative selection markers.

Notably, the present invention is the first demonstration of a TAT-derived peptide functioning as a periplasmic targeting sequence. Purification of fusion polypeptides of heterologously expressed proteins-of-interest from  
25 bacterial cells often result in poor yields, improperly folded or unfolded proteins, and proteins improperly processed, or impossible to purify from bacterial cell cytoplasm. The present invention circumvents these difficulties by providing a means to direct these heterologous proteins-of-interest to the periplasmic space, where proper folding often occurs as a function of the  
30 oxidizing environment of the intracellular compartment. Isolation and

purification of the protein is more readily accomplished as few bacterial proteins are located within the periplasmic space. Finally, addition of the TAT-derived peptide resulted in markedly higher production of the protein-of-interest following induction, and hence provides a superior means for isolation and purification of fusion-polypeptides containing a protein-of-interest or the protein-of-interest itself in a mature form.

In another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide harboring an intact polylinker cloning sequence operably linked to the TAT-derived peptide and a third polynucleotide harboring a prokaryotic promoter, operably linked to the TAT-derived peptide. This expression construct provides a means therefore of producing a fusion polypeptide containing a protein-of-interest subcloned within the polylinker cloning sequence, that when expressed will be targeted to the bacterial periplasm. In preferred embodiments of the present invention a reporter gene may be included in the construct as well, as are positive and/or negative selection markers.

As used herein the phrase "intact polylinker cloning sequence" refers to a non-interrupted polylinker cloning sequence which includes at least 50, 60, 70 or at least 80 nucleotides and harbors at least 2, preferably, at least 5, at least 10, at least 15 or at least 20 unique restriction endonuclease recognition sequences, some of which may be overlapping in sequence, at least one of which, preferably 2-10 of which are recognized by 6 cutter restriction endonucleases and at least one is recognized by a 8 cutter restriction endonuclease.

Similarly, mammalian secreted proteins-of-interest may be produced and targeted to the bacterial periplasm by incorporation within a bacterial expression construct encoding a TAT-derived peptide. Accordingly, in yet another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-

derived peptide, a second polynucleotide encoding a mammalian secreted protein-of-interest in frame with the TAT-derived peptide, and a third polynucleotide harboring a reporter gene, being operably linked to said second polynucleotide. In preferred embodiments of this invention a bacterial protease recognition sequence may be encoded as well, in order to provide a means of expressing a mature and processed mammalian secreted protein-of-interest. Positive and negative selection markers may be incorporated as indicated above.

Similarly, non-nuclear mammalian proteins-of-interest may be produced and targeted to the bacterial periplasm by incorporation within a bacterial expression construct encoding a TAT-derived peptide. Accordingly, in yet another aspect of the present invention there is provided a nucleic acid expression construct comprising a first polynucleotide encoding a TAT-derived peptide, a second polynucleotide encoding a non-nuclear mammalian protein-of-interest in frame with the TAT-derived peptide, and a third polynucleotide harboring a reporter gene, being operably linked to said second polynucleotide. In additional preferred embodiments of this invention a bacterial protease recognition sequence may be encoded, in order to provide a means of expressing a mature and processed mammalian non-nuclear protein-of-interest. Positive and negative selection markers may be incorporated as well.

It will be appreciated in this respect that the prior art fails to teach, suggest or create motivation for the production of such secreted and/or other non-nuclear mammalian (e.g., cytoplasmic) proteins using TAT-derived peptides linked thereto, because prior art production of TAT-derived protein-of-interest proteins was aimed at testing the ability of the TAT-derived peptide to transport the protein-of-interest to the nucleus of mammalian cells.

In still another preferred embodiment, in the expression constructs described hereinabove, the TAT-derived peptide is N-terminal to the protein-of-interest, when expressed.

In still other preferred embodiments, in the expression constructs described hereinabove, the TAT-derived peptide is N-terminal to the signal sequence.

In addition to expression constructs, the present invention provides for  
5 prokaryotic cells containing the constructs described herein and designed for expression of a protein-of-interest.

Accordingly, in another aspect of the present invention there is provided a prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a signal sequence and a protein-of-interest, wherein the  
10 TAT-derived peptide serves for transport of the fusion polypeptide to a periplasm of the prokaryotic cell and the signal sequence facilitates processing of the fusion polypeptide to yield a mature protein consisting essentially of the protein-of-interest and lacking the TAT-derived peptide and signal sequence.

Similarly, in yet another aspect of the present invention there is  
15 provided a prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a protease cleavage recognition sequence and a protein-of-interest, wherein the protease cleavage recognition sequence is positioned between the TAT-derived peptide and protein-of-interest, whereby the TAT-derived peptide serves for transport of the fusion  
20 polypeptide to a periplasm of the prokaryotic cell.

References to a protein-of-interest (either mammalian or not), mammalian secreted protein-of-interest and/or mammalian non-nuclear protein-of-interest may include proteins, such as, but not limited to, an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a  
25 glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral  
30 factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory



factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

Prokaryotic cells are thus utilized for expressing any and all of the expression constructs listed herein, and are a means of producing the recombinant proteins-of-interest.

in a preferred embodiment of the present invention, the prokaryotic cells utilized in any of these applications are of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

Depending on the strain selected, the expression construct will contain an appropriate origin-of-replication sequence for the respective prokaryotic strain employed for expression, selection markers and sequences encoding reporter polypeptides, and promoter or enhancer sequences operably linked to ensure the functioning of an expression control sequence.

It is to be understood that the expression constructs, bacterial strains and conditions for introducing the constructs can be optimized empirically and hence provide a readily accessible system for heterologous expression of proteins-of-interest in prokaryotes.

According to a still further aspect of the present invention there is provided a kit comprising any of the expression constructs described herein and optionally enzymes, substrates and/or reagents for expression, verification and utilization of the expression constructs of the invention.

According to a preferred embodiment of the present invention, the kit further comprises cells into which the expression construct can be transformed or transfected.

According to still another preferred embodiment, the kit further  
5 comprises cells competent for genetic manipulation.

According to still another preferred embodiment, the kit further comprises oligonucleotide primers for the amplification, purification and subcloning of specific sequences of interest within the polylinker cloning sequence within the prokaryotic expression construct.

10

Additional objects, advantages, and novel features of the present invention will become apparent to one ordinarily skilled in the art upon examination of the following examples, which are not intended to be limiting. Additionally, each of the various embodiments and aspects of the present  
15 invention as delineated hereinabove and as claimed in the claims section below finds experimental support in the following examples.

### EXAMPLES

Reference is now made to the following examples, which together with  
20 the above descriptions illustrate the invention in a non-limiting fashion.

Generally, the nomenclature used herein and the laboratory procedures utilized in the present invention include molecular, biochemical, microbiological and recombinant DNA techniques. Such techniques are thoroughly explained in the literature. See, for example, "Molecular Cloning:  
25 A laboratory Manual" Sambrook et al., (1989); "Current Protocols in Molecular Biology" Volumes I-III Ausubel, R. M., ed. (1994); Ausubel et al., "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, New York (1988); Watson et al., "Recombinant DNA",  
30 Scientific American Books, New York; Birren et al. (eds) "Genome Analysis:

A Laboratory Manual Series", Vols. 1-4, Cold Spring Harbor Laboratory Press, New York (1998); methodologies as set forth in U.S. Pat. Nos. 4,666,828; 4,683,202; 4,801,531; 5,192,659 and 5,272,057; "Cell Biology: A Laboratory Handbook", Volumes I-III Cellis, J. E., ed. (1994); "Current Protocols in Immunology" Volumes I-III Coligan J. E., ed. (1994); Stites et al. (eds), "Basic and Clinical Immunology" (8th Edition), Appleton & Lange, Norwalk, CT (1994); Mishell and Shiigi (eds), "Selected Methods in Cellular Immunology", W. H. Freeman and Co., New York (1980); available immunoassays are extensively described in the patent and scientific literature, see, for example, U.S. Pat. Nos. 3,791,932; 3,839,153; 3,850,752; 3,850,578; 3,853,987; 3,867,517; 3,879,262; 3,901,654; 3,935,074; 3,984,533; 3,996,345; 4,034,074; 4,098,876; 4,879,219; 5,011,771 and 5,281,521; "Oligonucleotide Synthesis" Gait, M. J., ed. (1984); "Nucleic Acid Hybridization" Hames, B. D., and Higgins S. J., eds. (1985); "Transcription and Translation" Hames, B. D., and Higgins S. J., eds. (1984); "Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) and "Methods in Enzymology" Vol. 1-317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak et al., "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996); all of which are incorporated by reference as if fully set forth herein. Other general references are provided throughout this document. The procedures therein are believed to be well known in the art and are provided for the convenience of the reader. All the information contained therein is incorporated herein by reference.

## EXAMPLE 1

## RECOMBINANT HUMAN GROWTH HORMONE EXPRESSION

*Materials and Experimental Methods*

5        *Isolation of the coding region for the mature polypeptide of human Growth Hormone:*

A 576 bp fragment of the mature human Growth Hormone cDNA (SEQ ID NO: 1) was amplified from Human Pituitary Gland cDNA (Clontech, Cat. No. 7173-1) by PCR using the specific sense primer: hGH.F – 5'-  
10 GGGCTATGCATTCCCAACCA TTCCGTTATCCAGGC-3' (SEQ ID NO: 2) and a specific antisense primer: hGH.R- 5'-ACCCGGATCCCTAGAAGCCACAGCTGCCCTCCACAG-3' (SEQ ID NO: 3). PCR conditions were: denaturation at 94 °C for 2 minutes, addition of heat stable DNA polymerase and additional denaturation at 94 °C for 40 seconds  
15 followed by annealing at 60 °C for 80 seconds. 35 cycles of the following steps were then carried out: elongation at 72 °C for 3 minutes; denaturation at 94 °C for 40 seconds; and annealing at 60 °C for 80 seconds. The last step in the amplification was elongation at 72 °C for 5 minutes.

The hGH.F primer introduced a *NsiI* site and an in frame ATG codon.  
20 The hGH.R primer introduced a *BamHI* site and a translation stop codon. The sequence of the PCR product was confirmed with sequence specific primers, using an automated DNA sequencer (Applied Biosystems, Model 373A).

*Construction of IPTG-induced expression constructs:*

An expression construct, harboring the *E. coli* heat-stable enterotoxin II  
25 signal peptide (Picken et al, "Nucleotide Sequence of the Gene for Heat-Stable Enterotoxin II of *Escherichia coli*", Infect. Immun. 42(1): 269-275, 1983), which enables targeting of recombinant products to the periplasm and cleavage of the amino terminal to liberate a mature mammalian hGH protein, was assembled as follows:

The DNA fragment encoded for the heat-stable enterotoxin STI signal peptide (designated as STB) was constructed by annealing oligonucleotides STB.F: TATGAAAAAG AATATCGCAT TTCTTCTTGC ATCTATGTTC GTTTTTTCTA TTGCTACAAA TGCCTATGCA TG (SEQ ID NO: 4) and STB.R: GATCCATGCA TAGGCATTG TAGCAATAGA AAAAACGAAC ATAGATGCAA GAAGAAATGC GATATTCTTT TTCA (SEQ ID NO: 5). The resulting double-stranded DNA fragment has a compatible *NdeI* overhang at the 5' end and an *NsiI* overhang at the 3' end that was generated following digestion with *NsiI* endonuclease. The PCR product of the mature hGH cDNA was then digested with *NsiI* and *BamHI*. A 3-piece ligation of the STB and the mature hGH cDNA into the *NdeI* and *BamHI* sites in the pET-24b plasmid (Novagen) resulted in the establishment of an expression construct, designated pSTB-hGH. This expression construct encoded an STB-hGH open reading frame of 214 amino acids (SEQ ID NO: 6).

A second expression construct, containing the TAT-derived sequence amino terminal to the mature hGH was constructed as follows: The TAT-derived sequence was generated by annealing of two oligonucleotides mTAT1.F: CATATGAAAG GCTATGGCCG CAAAAAACGT CGCCAGCGTC GCCGTGGTGC A (SEQ ID NO: 7) and mTAT1.R: CCACGGCGAC GCTGGCGACG TTTTGTGCGG CCATAGCCTT TCATATG (SEQ ID NO: 8). This resulted in generation of the mTAT1 sequence with an *NdeI* overhang at the 5' end and *NsiI* site at the 3' end. The mTAT1 was then ligated with mature hGH removed from the plasmid pSTB-hGH with *NsiI*-*BamHI* endonuclease. The ligation product was PCR amplified with a sense primer mTAThGH.F: AACATATGAA AGGCTATGGC CGCAA (SEQ ID NO: 9) and an antisense primer mTAThGH.R: AAAGGATCCA TTAGAAGCCA CAGCTGCCCT C (SEQ ID NO: 10). The PCR product was digested with *NdeI* and *BamHI* endonucleases and ligated into the corresponding sites in the pET-24b plasmid (Novagen). The

expression construct obtained, designated pTAT-hGH, and encoded an TAT-hGH open reading frame of 207 amino acids (SEQ ID NO: 11).

An additional expression construct containing both the TAT-derived sequence and the STB sequence amino terminal to the mature hGH cDNA was assembled as follows: Two synthetic oligonucleotides mTAT2.F: 5 CATATGAAAG GCTATGGCCG CAAAAAACGT CGCCAGCGTC GCCGTGGCGC A (SEQ ID NO: 12) and mTAT2.R: CCACGGCGAC GCTGGCGACG TTTTGTGCGG CCATAGCCTT TCATATG (SEQ ID NO:13) were annealed, generating the mTAT2 sequence with an *NdeI* overhang at the 5' end and at the 3' end a site compatible for ligation with the 5' of cSTB fragment. The cSTB fragment was assembled by annealing oligonucleotides cSTB.F: TTTCTTCTTG CATCTATGTT CGTTTTTCT ATTGCTACAA ATGCCTATGC A (SEQ ID NO: 14) and cSTB.R: TAGGCATTTG TAGCAATAGA AAAAACGAAC ATAGATGCAA 15 GAAGAAATGC G (SEQ ID NO:15). The cSTB fragment has a 5' end compatible for ligation with the 3' end of mTAT2 and 3' end, which is *NsiI* compatible. The mTAT2 and cSTB fragments were then ligated to an *NsiI*-*BamHI* fragment of the mature hGH cDNA obtained from plasmid pSTB-hGH. The ligation product was PCR amplified with a sense primer 20 mTAThGH.F (SEQ ID NO: 9) and antisense primer mTAThGH.R (SEQ ID NO: 10). The PCR product was digested with *NdeI* and *BamHI* endonucleases and ligated into the corresponding sites in pET24b (Novagen). The resulting expression construct, designated pTAT-STB-hGH, encoded a TAT-STB-hGH open reading frame of 224 amino acids (SEQ ID NO: 16). A schematic 25 representation of the expression construct pTAT-STB-hGH is shown in Figure 1.

#### *Construction of a heat-inducible expression construct:*

Construction of the heat inducible expression constructs hGHSP/pACYC184 and hGHTSP/pACYC184 was a multiple step process. An 30 expression cassette containing the following components was assembled

(Figure 2a):  $\lambda$  gt11 thermo-labile repressor under the control of a synthetic constitutive promoter; the phase  $\lambda$  pL promoter; a multiple cloning site (MCS); and a transcriptional termination site (TT). Transcription from the constitutive promoter controlling the  $\lambda$  gt11 thermo-labile repressor is in opposite direction to transcription from the  $\lambda$  pL promoter. Assembly of the cassette components resulted in the production of a DNA fragment (SEQ ID NO: 17), flanked by unique *KpnI* and *SacI* sites (Figure 2A). The cassette was initially ligated into the corresponding *SacI-KpnI* sites in the vector pBluescript II SK- (Stratagene, GeneBank Accession No. X52330), generating the vector pV1Rep1. DNA fragments containing the signal peptide, hGH cDNA and ribosomal binding site (RBS) sequences, were then removed using *XbaI-BamHI* digestion from plasmids pTAT-STB-hGH and pSTB-hGH, described above. These fragments were ligated to an *XbaI-BamHI* digested pV1Rep1 vector. The expression constructs generated are hGHTSP/pV1Rep1 and hGHSP/pV1Rep1 with and without the TAT-derived sequence immediately following the initial Methionine translation start site, respectively.

The entire expression cassette including the regulatory elements and the hGH sequence were removed from vectors hGHTSP/pV1Rep1 and hGHSP/pV1Rep1 by *SacI-KpnI* digestion. The *SacI-KpnI* fragments were blunt-ended using T7 DNA polymerase and ligated to a blunt-ended *AhdI-XmnI* fragment from the plasmid pACYC184 (New-England BioLabs, GeneBank Accession No. X06403). The ligations generated expression constructs hGHTSP/pACYC184 and hGHSP/pACYC184, with and without the TAT-derived sequence, respectively. Schematic representation of the expression construct hGHTSP/pACYC184 is shown in Figure 2B.

***Protein expression using an IPTG induction system:***

The three constructs pSTB-hGH, pTAT-hGH and pTAT-STB-hGH, were used to transform competent cells of *E. coli* strain BL21(DE3) (Stratagene). Single colonies were selected. Bacteria, harboring these plasmids, were grown in medium containing 10 grams/liter (g/L) tryptone, 5

g/L yeast extract and 10 g/L sodium chloride supplemented with 30 µg/ml of kanamycin (LB-kan). Following overnight growth, cultures were backdiluted 1:20 in fresh LB medium (10 g/L tryptone, 5 g/L yeast extract and 10 g/L sodium chloride) and grown for 2.5 hours at 30 °C. Protein expression was induced with media supplementation with 1 mM IPTG (isopropyl-beta-D-1-thiogalactopyranoside) for an additional 2.5 hours. One milliliter of bacteria was centrifuged (18,500 g, 5 minutes] and the pellet resuspended at a concentration yielding an optical density (OD) of 5 at 600 nm. The bacteria was subjected to osmotic shock by incubation with a solution containing 20 mM Tris-HCl at pH 8, 2.5 mM EDTA (ethylenediaminetetraacetic acid, pH 8.0) and 20 % (w/v) sucrose (osmotic shock solution No. 1), and incubated for 10 minutes at 4 °C. Following centrifugation (18,500 g, 5 minutes) the cell pellet was resuspended in the same volume as above with a solution containing 20 mM Tris-HCl, pH 8.0 and 2.5 mM EDTA (pH 8.0) (osmotic shock solution No. 2) and incubated for 10 minutes at 4 °C. After centrifugation the supernatant contained the periplasmic proteins released from the osmotic-shocked bacteria, and the pellet contained the cytoplasmic proteins. Equal volumes of each of the samples were diluted in SDS-PAGE loading buffer and separated on a 4-20% denaturing gel (Novex). For Western blot analysis the proteins were transferred onto a nitrocellulose membrane, incubated with goat anti-hGH antibodies (Santa Cruz, Cat. No. SC-10365) and detected with an ECL kit (Amersham Pharmacia).

### *Experimental Results*

#### *Expression of the amino terminal TAT fusion polypeptide in bacteria:*

Several bacterial clones harboring the constructs pSTB-hGH and pTAT-STB-hGH were cultured and induced to express the recombinant proteins. Protein samples were collected for analysis before and following induction, from the periplasmic compartment of osmotic-shocked bacteria and from the remaining cytoplasmic compartment. The Coomassie Blue stained gel



(Figure 3) revealed pronounced cytoplasmic accumulation of TAT-STB-hGH polypeptide, which was much higher than the cytoplasmic accumulation of the STB-hGH polypeptide. The periplasmic mature hGH had an apparent molecular weight lower than the cytoplasmic hGH due to processing and removal of the signal peptide following transport to the periplasm. Higher levels of the mature periplasmic hGH were found in bacteria harboring the pTAT-STB-hGH construct, as opposed to bacteria harboring the STB-hGH alone construct (Figure 3). This difference was clearly evident in the antibody-probed Western blot (Figure 4, lower panel), where the level of mature hGH was highest in bacteria harboring the pTAT-STB-hGH construct (~22 kDa), following induction. This Figure also shows that the TAT-hGH fusion polypeptide (lacking the heat-stable enterotoxin II signal peptide) efficiently translocated into the periplasmic compartment as a non-processed fusion polypeptide (~24 kDa).

Thus TAT-derived sequences inserted into a periplasmic-targeting signal peptide enabled the transport of the fusion polypeptide to the periplasm, correct processing of the synthetic signal sequence, and enhanced the level of accumulation of the mature polypeptide in the periplasm.

## EXAMPLE 2

### RECOMBINANT TAT-STB-hGH EXPRESSION RESULTS IN PROPER PROTEIN PROCESSING

#### *Material and Experimental Methods*

#### *Expression, isolation and determination of the N-terminal protein sequence of mature recombinant hGH*

The heat inducible expression construct hGHTSP/pACYC184 was used to transform *E. coli* strain MM294 (ATCC 33625) establishing the hGH expression clone. Transformation of the expression construct to *E. coli* was carried out by electroporation using the BioRad Micro Pulser Electroporator

(Cat. No. 165-2100). Transformation was performed according to the manufacturer recommendations. Transformants were plated on LB agar plates supplemented with 12 µg/ml tetracycline, and incubated overnight at 30 °C. A single colony was inoculated into broth containing 5 ml LB-tet medium (10 g/L tryptone, 5 g/L yeast extract and 10 g/L sodium chloride supplemented with 12 µg/ml of tetracycline) and incubated at 30 °C for roughly 14 hours. The culture was then backdiluted to an OD of 0.1 (600 nm) in fresh LB medium (without tetracycline). The bacterial culture was grown at 30 °C for 2 hours and protein expression was induced by elevation of the culture temperature to 42 °C for 6 hours. Cells were harvested and periplasmic proteins were extracted using the following method: Cells were brought to an OD of 5 at 600 nm. Osmotic shock was induced with osmotic shock solution No. 1 (detailed in Example 1 above). Cells were incubated on ice for 10 minutes followed by additional centrifugation (18,500 g, 5 minutes). The cell pellet was resuspended in the same volume as above with osmotic shock solution No. 2 (detailed in Example 1). Cells were then incubated on ice for 10 minutes followed by centrifugation (18,500 g, 5 minutes). The supernatant obtained (periplasmic fraction) was concentrated using the YM-10 Centricon (Millipore, Cat. No. 4321), following manufacturer's instructions and kept overnight at 4 °C.

One ml of the above concentrate was acidified by the addition of 15 µl glacial acetic acid. The acidified concentrate was filtered through a 0.45 micron PVDF syringe driven filter unit (Millex-HV, Millipore, Cat. No.: SLHV R04 NL) and 250 - 900 µl were injected on a Waters Delta Prep HPLC system (Delta Prep 4000, Preparative Chromatography System) fitted with a Vydac C4, 300 A, reverse phase column (4.6 x 250 mm). The mobile phase was: A - 0.1 % trifluoroacetic acid in water, B - 0.1% trifluoroacetic acid in acetonitrile. The flow rate was 1 ml/minute. The gradient used was 20-80 % of B in 60 minutes. Detection was done at 280 nm. The standard (Genotropin,

Pharmacia-Upjohn, Sweden) eluted under these conditions with a retention time of 24.5 minutes.

Determination of the amino-terminal sequence sample eluted from the RP-HPLC column was carried out following electrophoresis of the sample on  
5 14 % SDS-PAGE (Novex) and transferred to a PVDF membrane (Immobilone-P, Millipore, IPVH00010). Amino terminal sequence was determined using an Applied Biosystems Procise Sequencer (Model 494).

### *Experimental Results*

#### 10 *Correct cleavage of the mature recombinant hGH polypeptide:*

Attempts were carried out to find whether incorporation of a TAT-derived sequence into the *E. coli* signal peptide does not interfere with correct processing of the signal peptide by the cell machinery. Periplasmic hGH isolated from *E. coli* MM294 harboring the expression construct  
15 hGHTSP/pACYC184, containing the TAT-derived sequence upstream to the heat-stable enterotoxin II signal peptide, was subjected to N-terminal amino acid sequencing. The results indicated a correct processing of the hGH. The 5 N-terminal amino acids detected, in the following order, are: Phenylalanine (F), Proline (P), Threonine (T), Isoleucine (I) and Proline (P). The amino acid  
20 sequence is identical to that of the mature hGH (DeNoto et al. "Human Growth Hormone DNA Sequence and mRNA Structure: Possible Alternative Splicing", Nucleic Acids Research, 9(15):3719-3730, 1981; Seeburg, "The Human Growth Hormone Gene family: Nucleotide Sequences Show Recent Divergence and Predict a New Polypeptide Hormone" DNA, 1(3):239-249,  
25 1982), indicating that TAT sequence incorporation enables transport and accumulation of the soluble fusion polypeptide in the periplasmic space, where processing occurs, providing the mature hGH product.

## EXAMPLE 3

TAT SEQUENCE INCLUSION ENHANCES PROKARYOTIC hGH  
PRODUCTION

5

*Material and Experimental Methods**Large-scale fermentation, recombinant hGH recovery:*

Five ml of starter cultures of *E. coli* MM294 containing expression constructs hGHSP/pACYC184 or hGHTSP/pACYC184, were grown in LB-tet medium (10 g/L tryptone, 5 g/L yeast extract and 10 g/L sodium chloride  
10 supplemented with 12 µg/ml of Tetracycline) at 30 °C for 8-10 hours. Cells from the starter medium were inoculated into 300 ml LB-tet medium and grown for an additional 13-16 hours at 30 °C. The 300 ml cultures were then inoculated into a 5 liter fermenter (BIOFLOW 3000, New-Brunswick, USA). The initial fermentation medium contained the following ingredients:  
15 21.9 mM potassium dibasic phosphate, 13.9 mM sodium monobasic phosphate, 29.6 mM potassium chloride, 55.7 mM ammonium sulfate, 5 mM sodium citrate, 14.7 mM magnesium sulfate, 1.11 % tryptone, 1.11 % yeast extract, 0.11 % glucose, 0.002 % ferric sulfate, 0.4 ml/L of antifoam and 0.5 mg/L tetracycline. A trace element solution (3.7 ml/5 L) was added, containing  
20 100 mM ferric sulfate and 30 mM of each of the following: zinc sulfate, cobalt chloride, sodium molybdate, copper sulfate, boric acid and manganese sulfate.

A pH of 7.2 was maintained throughout the process by the addition of H<sub>2</sub>SO<sub>4</sub> solution and ammonium hydroxide. 50 % (w/v) of a glucose solution was fed to the fermenter from initiation of the process. Adjustment of glucose  
25 feeding was done during the process in order to maintain a glucose level of below 4 g/L. Dissolved oxygen was measured by an on-line oxygen electrode and was set to 30%. The dissolved oxygen setting was maintained by increasing the agitation, airflow and oxygen supplementation during the process.

Cell growth was performed at 30 °C for 5.5 hours, followed by increasing the culture temperature to 42° C over a period of 30 minutes. Induction of hGH expression was performed at 42° C for 6 hours. Bacterial cells were harvested via centrifugation (6,200 g, 15 minutes at 4°C) and the  
5 cell pellet was stored at -20° C.

Extraction of total cell proteins was performed on 8 grams of wet cells: cells were frozen at -20 °C, thawed at 25 °C, resuspended in 200 ml buffer containing 10 mM Tris-HCl pH 8.0 and 1 mM EDTA pH 8.0 and homogenized (Ultra-Turrax homogenizer, T50 basic, IKA-WERKE,  
10 Germany). 12.5 mg of lysozyme (Roche, Cat. No. 107255, 135,000 U/mg) was added, cells were incubated for 30 minutes at 25 °C and then briefly homogenized, as above. Cells were then passed twice through a Gaulin Lab 1000 (APV, Denmark) homogenizer using a pressure of 600-800 bars. The supernatant, defined as the total cell extract, was removed following  
15 centrifugation at 9,000 g for 30 minutes at 4 °C and aliquots were stored at -70 °C.

Determination of total soluble hGH in the total cell extract was performed using an hGH ELISA kit (Roche, Cat. No. 1585878).

20

### *Experimental Results*

#### *Large scale expression of amino terminal TAT-fusion polypeptides in bacteria:*

In order to determine the effect of TAT-derived sequences on large-scale production of recombinant hGH, two *E. coli* MM294 clones harboring  
25 expression constructs hGHSP/pACYC184 and hGHTSP/pACYC184 were used in large-scale experiments. Surprisingly, the presence of the TAT-derived sequence in *E. coli* clones containing the expression construct hGHTSP/pACYC184, resulted in higher cell density following induction at the permissive temperature, as compared to *E. coli* clones harboring the expression  
30 construct lacking the TAT-derived sequence (Figure 5). Whereas decline in

cell growth was observed following the prolonged induction stage in hGHSP/pACYC184 expressing strains, continued growth was maintained for more than 12 hours, in cells expressing the TAT-derived sequence (Figure 5). Differences in growth pattern were not apparent during the growth phase at 30 °C, but rather following induction. A total cell mass of roughly 78 gram wet cells/ L, as compared to roughly 53 gram wet cells/ L was obtained for *E. coli* clones harboring the expression construct hGHTSP/pACYC184 and hGHSP/pACYC184, respectively.

Analysis by ELISA of the total soluble recombinant hGH formed following 6 hours induction at the permissive temperature indicated that twice as much volumetric productivity was obtained by clones harboring the TAT-derived sequence as compared to those without (roughly 107 mg hGH /L versus 53 mg hGH/L, respectively). Thus, incorporation of TAT sequences inserted into a periplasmic-targeting signal peptide within prokaryotic expression constructs in a prokaryotic clone, expressing recombinant hGH, enabled accumulation of significantly larger cell mass during protein induction than that which accumulated in the same *E. coli* clone, lacking the TAT-derived sequence. TAT incorporation enhanced mature protein expression, by increasing bacterial cell mass.

20

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable subcombination.

25

Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art. Accordingly, it is

30

intended to embrace all such alternatives, modifications and variations that fall within the spirit and broad scope of the appended claims. All publications, patents, patent applications and sequences identified by their accession numbers mentioned in this specification are herein incorporated in their  
5 entirety by reference into the specification, to the same extent as if each individual publication, patent, patent application or sequence identified by their accession number was specifically and individually indicated to be incorporated herein by reference. In addition, citation or identification of any reference in this application shall not be construed as an admission that such  
10 reference is available as prior art to the present invention.

## WHAT IS CLAIMED IS:

1. A method of producing a protein-of-interest in, and purifying the protein-of-interest from, bacteria, the method comprising:
  - (a) introducing into the bacteria an expression construct encoding a fusion polypeptide which comprises a TAT-derived peptide, a signal sequence and the protein-of-interest, said TAT-derived peptide serving for transport of the fusion polypeptide from a cytoplasm to a periplasm of the bacteria, said signal sequence facilitating processing the fusion polypeptide to a mature protein, said mature protein consisting essentially of said protein-of-interest and substantially lacking said TAT-derived peptide and said signal sequence, in said periplasm; and
  - (b) purifying said mature protein substantially exclusively from said periplasm.
2. The method of claim 1, wherein said bacteria is a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.
3. The method of claim 1, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).
4. The method of claim 1, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.



5. The method of claim 1, wherein said expression construct encoding said fusion polypeptide which comprises said TAT-derived peptide and said protein-of-interest is engineered such that said TAT-derived peptide is N-terminal to said protein-of-interest.

6. The method of claim 1, wherein said expression construct encoding said fusion polypeptide which comprises said TAT-derived peptide and said signal sequence is engineered such that said TAT-derived peptide is N-terminal to said signal sequence.

7. The method of claim 1, wherein said signal sequence comprises a positively charged amino-terminus, a hydrophobic central region, and a neutral but polar carboxy-terminus.

8. The method of claim 1, further comprising a reporter gene.

9. The method of claim 8, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

10. The method of claim 1, wherein said expression construct further comprises a promoter operably linked to the polynucleotides encoding said fusion polypeptide.

11. the expression construct of claim 10, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

12. The method of claim 1, wherein said protein-of-interest is

selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

13. The method of claim 1, wherein purifying said protein-of interest substantially exclusively from said bacterial periplasm comprises:

- (a) isolating the bacterial periplasmic compartment from other subcellular compartments;
- (b) lysing the bacterial periplasmic compartment; and
- (c) purifying the protein-of-interest.

14. A method of producing a fusion polypeptide in, and purifying the fusion polypeptide from, bacteria, the method comprising:

- (a) introducing into the bacteria an expression construct encoding the fusion polypeptide which comprises a TAT-derived peptide, and a protein-of-interest, said TAT-derived peptide serving for transport of the fusion polypeptide from a cytoplasm to a periplasm of said bacteria; and
- (b) purifying the fusion polypeptide substantially exclusively from

said periplasm.

15. The method of claim 14, wherein said fusion polypeptide further comprises a protease cleavage recognition sequence positioned between said TAT-derived peptide and said protein-of-interest, the method further comprising cleaving said fusion polypeptide with a protease specific to said protease cleavage recognition sequence.

16. The method of claim 14, wherein said bacteria is a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

17. The method of claim 14, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

18. The method of claim 14, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

19. The method of claim 14, wherein said expression construct encoding said fusion polypeptide which comprises said TAT-derived peptide and said protein-of-interest is engineered such that said TAT-derived peptide is N-terminal to said protein-of-interest.

20. The method of claim 14, further comprising a reporter gene.

21. The expression construct of claim 20, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

22. The method of claim 14, wherein said expression construct further comprises a promoter operably linked to the polynucleotides encoding said fusion polypeptide.

23. The expression construct of claim 22, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

24. The method of claim 14, wherein said protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

25. The method of claim 14, wherein purifying said protein-of interest substantially exclusively from said bacterial periplasm comprises

- (a) isolating the bacterial periplasmic compartment from other subcellular compartments;
- (b) lysing the bacterial periplasmic compartment; and
- (c) purifying the polypeptide.

26. A nucleic acid expression construct comprising:

- (a) a first polynucleotide encoding a TAT-derived peptide;
- (b) a second polynucleotide harboring an intact polylinker cloning sequence, said intact polylinker cloning sequence being operably linked to said first polynucleotide; and
- (c) a third polynucleotide harboring a prokaryotic promoter, being operably linked to said first polynucleotide.

27. The expression construct of claim 26, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

28. The expression construct of claim 26, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

29. The expression construct of claim 26, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

30. The expression construct of claim 29, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

31. The expression construct of claim 29, further comprising a reporter gene.

32. The expression construct of claim 31, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

33. The expression construct of claim 31, further comprising a polynucleotide encoding a protease cleavage recognition sequence in frame with said TAT-derived peptide.

34. The expression construct of claim 31, further comprising a polynucleotide encoding a positive or a negative selection marker.

35. A nucleic acid expression construct comprising:

- (a) a first polynucleotide encoding a TAT-derived peptide;
- (b) a second polynucleotide encoding a signal sequence in frame with said TAT-derived peptide;
- (c) a third polynucleotide harboring a polylinker cloning sequence, being operably linked to said second polynucleotide; and
- (d) a fourth polynucleotide harboring a prokaryotic promoter, being operably linked to said first polynucleotide.

36. The expression construct of claim 35, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

37. The expression construct of claim 35, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

38. The expression construct of claim 35, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

39. The expression construct of claim 35, wherein said signal sequence comprises a positively charged amino-terminus, a hydrophobic central region, and a neutral but polar carboxy-terminus.

40. The expression construct of claim 35, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

41. The expression construct of claim 35, further comprising a reporter gene.

42. The expression construct of claim 35, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

43. The nucleic acid construct of claim 35, further comprising a polynucleotide encoding a positive or a negative selection marker.

44. A nucleic acid expression construct comprising:

(a) a first polynucleotide encoding a TAT-derived peptide;

- (b) a second polynucleotide encoding a protease cleavage recognition sequence in frame with said TAT-derived peptide;
- (c) a third polynucleotide encoding a protein-of-interest in frame with said protease cleavage recognition sequence; and
- (d) a fourth polynucleotide harboring a prokaryotic promoter, being operably linked to said first polynucleotide.

45. The expression construct of claim 44, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

46. The expression construct of claim 44, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

47. The expression construct of claim 44, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

48. The expression construct of claim 44, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

49. The expression construct of claim 44, further comprising a reporter gene.



50. The expression construct of claim 44, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

51. The nucleic acid construct of claim 44, further comprising a polynucleotide encoding a positive or a negative selection marker.

52. The expression construct of claim 44, wherein said protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

53. A nucleic acid expression construct comprising:

- (a) a first polynucleotide encoding a TAT-derived peptide,
- (b) a second polynucleotide encoding a signal sequence in frame with said TAT-derived peptide;
- (c) a third polynucleotide encoding a protein-of-interest in frame with said signal sequence; and

57

(d) a fourth polynucleotide harboring a prokaryotic promoter, being operably linked to said first polynucleotide.

54. The expression construct of claim 53, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

55. The expression construct of claim 53, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

56. The expression construct of claim 53, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

57. The expression construct of claim 53, wherein said signal sequence comprises a positively charged amino-terminus, a hydrophobic central region, and a neutral but polar carboxy-terminus.

58. The expression construct of claim 53, wherein said promoter is selected from the group consisting of constitutive and/or inducible prokaryotic promoter.

59. The expression construct of claim 53, further comprising a reporter gene.

60. The expression construct of claim 53, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

61. The nucleic acid construct of claim 53, further comprising a polynucleotide encoding a positive or a negative selection marker.

62. The expression construct of claim 53, wherein said protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

63. A nucleic acid expression construct comprising:

- (a) a first polynucleotide encoding a TAT-derived peptide,
- (b) a second polynucleotide encoding a protein-of-interest in frame with said TAT-derived peptide, said protein-of-interest is a mammalian secreted protein; and
- (c) a third polynucleotide harboring a reporter gene, being operably

linked to said second polynucleotide.

64. The expression construct of claim 63, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

65. The expression construct of claim 63, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

66. The expression construct of claim 63, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

67. The expression construct of claim 63, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

68. The nucleic acid construct of claim 63, further comprising a polynucleotide encoding a positive or a negative selection marker.

69. The expression construct of claim 63, wherein said mammalian secreted protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450

enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, and a substance P.

70. A nucleic acid expression construct comprising:

- (a) a first polynucleotide encoding a TAT-derived peptide,
- (b) a second polynucleotide encoding a protein-of-interest in frame with said TAT-derived peptide, said protein-of-interest is a mammalian, non-nuclear, protein; and
- (c) a third polynucleotide harboring a reporter gene, being operably linked to said second polynucleotide.

71. The expression construct of claim 70, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

72. The expression construct of claim 70, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

73. The expression construct of claim 70, engineered such that said first polynucleotide encoding said TAT-derived peptide is upstream of said second polynucleotide.

74. The expression construct of claim 70, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein.

75. The nucleic acid construct of claim 70, further comprising a polynucleotide encoding a positive or a negative selection marker.

76. The expression construct of claim 70, wherein said mammalian non-nuclear protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, and a substance P.

77. A kit, comprising the expression construct of claim 26.

78. The kit of claim 77, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

79. A kit, comprising the expression construct of claim 35.

80. The kit of claim 79, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

81. A kit, comprising the expression construct of claim 44.

82. The kit of claim 81, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

83. A kit, comprising the expression construct of claim 53.

84. The kit of claim 83, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

85. A kit, comprising the expression construct of claim 63.

86. The kit of claim 85, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

87. A kit, comprising the expression construct of claim 70.

88. The kit of claim 87, further comprising enzymes, substrates and/or reagents for expression, verification and utilization of said expression construct.

89. An assay of determining whether a TAT-derived peptide is an effective periplasmic targeting sequence, the assay comprising:

(a) introducing into bacteria an expression construct encoding a fusion polypeptide comprising the TAT-derived peptide and a reporter protein; and

(b) determining to what extent said fusion polypeptide accumulates within the periplasm, thereby determining whether the TAT-derived peptide is an effective periplasmic targeting sequence.

90. The assay of claim 89, wherein said TAT-derived peptide is derived from a virus selected from the group consisting of HIV 1, HIV-2, equine infectious anemia virus, simian immunodeficiency virus (SIV), bovine immunodeficiency virus (BIV), feline immunodeficiency virus (FIV), maedi-visna virus (MVV) and caprine arthritis-encephalitis-virus.

91. The assay of claim 89, wherein said reporter gene is selected from the group consisting of  $\beta$ -galactosidase, chloramphenicol acetyl transferase, luciferase and a fluorescent protein

92. The assay of claim 89, wherein determining said extent said fusion polypeptide accumulates within said periplasm comprises using at least one of the following assays: subcellular fractionation, column chromatography, Western blot analysis, HPLC, mass spectroscopy, GLC, immunocytochemistry and immunoelectron microscopy.

93. A prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a signal sequence and a protein-of-interest, wherein said TAT-derived peptide serves for transport of the fusion polypeptide to a periplasm of the prokaryotic cell and said signal sequence facilitates processing of the fusion polypeptide to yield a mature protein consisting essentially of said protein-of-interest and lacking said TAT-derived peptide and said signal sequence.



94. The prokaryotic cell of claim 93, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

95. The prokaryotic cell of claim 93, wherein said signal sequence comprises a positively charged amino-terminus, a hydrophobic central region, and a neutral but polar carboxy-terminus.

96. The prokaryotic cell of claim 93, wherein said protein-of-interest is selected from the group consisting of an insulin, an amylase, a protease, a lipase, a heparinase, a kinase, a phosphatase, a glycosyl transferase, a trypsinogen, a chymotrypsinogen, a carboxypeptidase, a hormone, a ribonuclease, a deoxyribonuclease, a triacylglycerol lipase, a phospholipase A2, an elastase, an amylase, a blood clotting factor, a UDP glucuronyl transferase, an ornithine transcarbamoylase, a cytochrome p450 enzyme, an adenosine deaminase, a serum thymic factor, a thymic humoral factor, thymopoietin, a growth hormone, a somatomedin, a costimulatory factor, an antibody, a colony stimulating factor, an erythropoietin, an epidermal growth factor, a hepatic erythropoietic factor (hepatopoietin), a liver-cell growth factor, an interleukin, an interferon, a negative growth factor, a fibroblast growth factor, a transforming growth factor of the  $\alpha$  family, a transforming growth factor of the  $\beta$  family, a gastrin, a secretin, a cholecystokinin, a somatostatin, a serotonin, a substance P, a transcription factor an avidin, a fluorescent protein and a streptavidin.

97. The prokaryotic cell of claim 93, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*,

*Francisella, Brucella, Legionella, Rickettsia, Coxiella, Haemophilus, Yersinia* and *Mycoplasma*.

98. A prokaryotic cell engineered to express a fusion polypeptide comprising a TAT-derived peptide, a protease cleavage recognition sequence and a protein-of-interest, wherein said protease cleavage recognition sequence is positioned between said TAT-derived peptide and said protein-of-interest, whereby said TAT-derived peptide serves for transport of the fusion polypeptide to a periplasm of the prokaryotic cell.

99. The prokaryotic cell of claim 98, wherein said TAT-derived peptide comprises the amino acid sequence YGRKKRRQRRR (SEQ ID NO:18).

100. The prokaryotic cell of claim 98, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

101. A prokaryotic cell engineered to express the construct of claim 44.

102. The prokaryotic cell of claim 101, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

103. A prokaryotic cell engineered to express the construct of claim 53.

104. The prokaryotic cell of claim 103, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

105. A prokaryotic cell engineered to express the construct of claim 63.

106. The prokaryotic cell of claim 105, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

107. A prokaryotic cell engineered to express the construct of claim 70.

108. The prokaryotic cell of claim 107, wherein said prokaryotic cell is of a strain of a species selected from the group consisting of *Escherichia*, *Streptococcus*, *Staphylococcus*, *Bacillus*, *Mycobacteria*, *Enterobacteriaceae*, *Vibrio*, *Campylobacter*, *Helicobacter*, *Neisseria*, *Pseudomonas*, *Listeria*, *Francisella*, *Brucella*, *Legionella*, *Rickettsia*, *Coxiella*, *Haemophilus*, *Yersinia* and *Mycoplasma*.

1/3

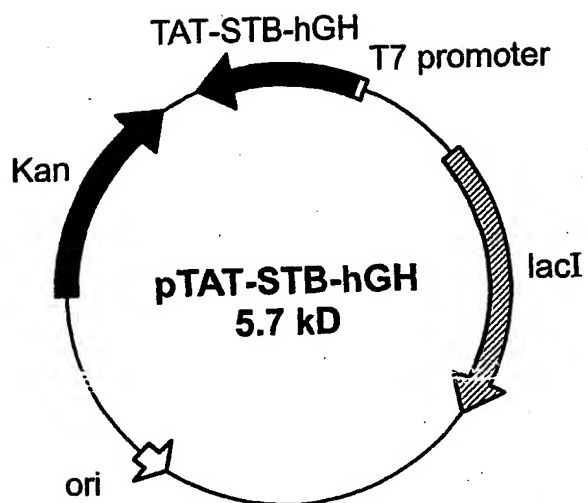


Fig. 1

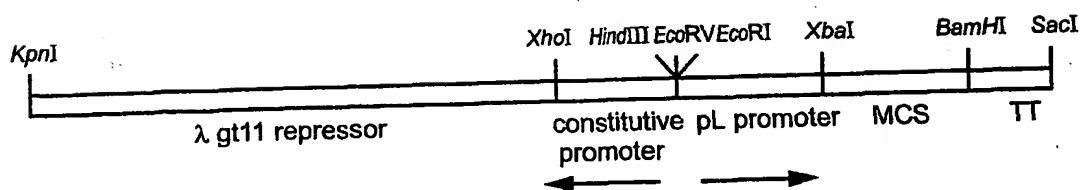


Fig. 2a

2/3

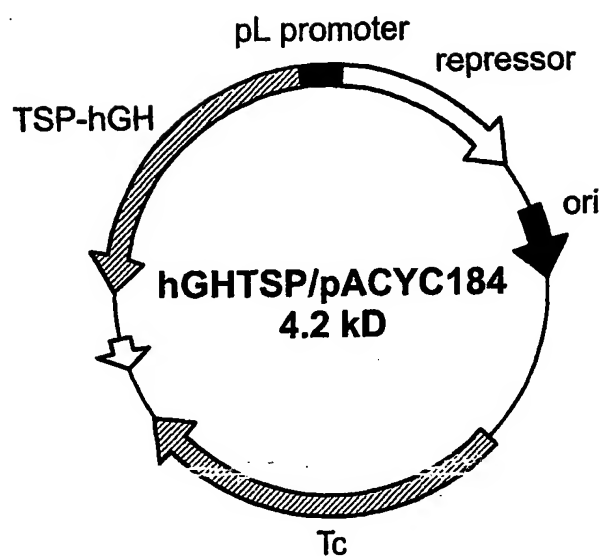


Fig. 2b

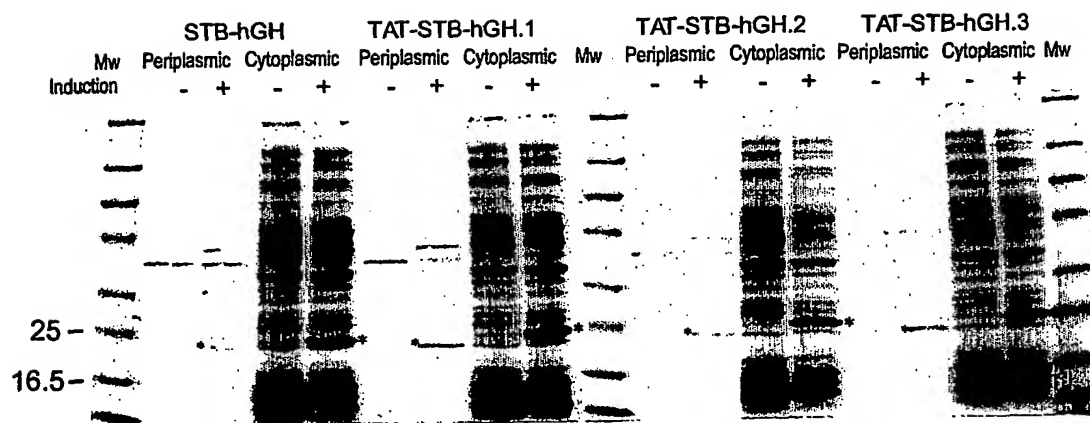


Fig. 3

3/3

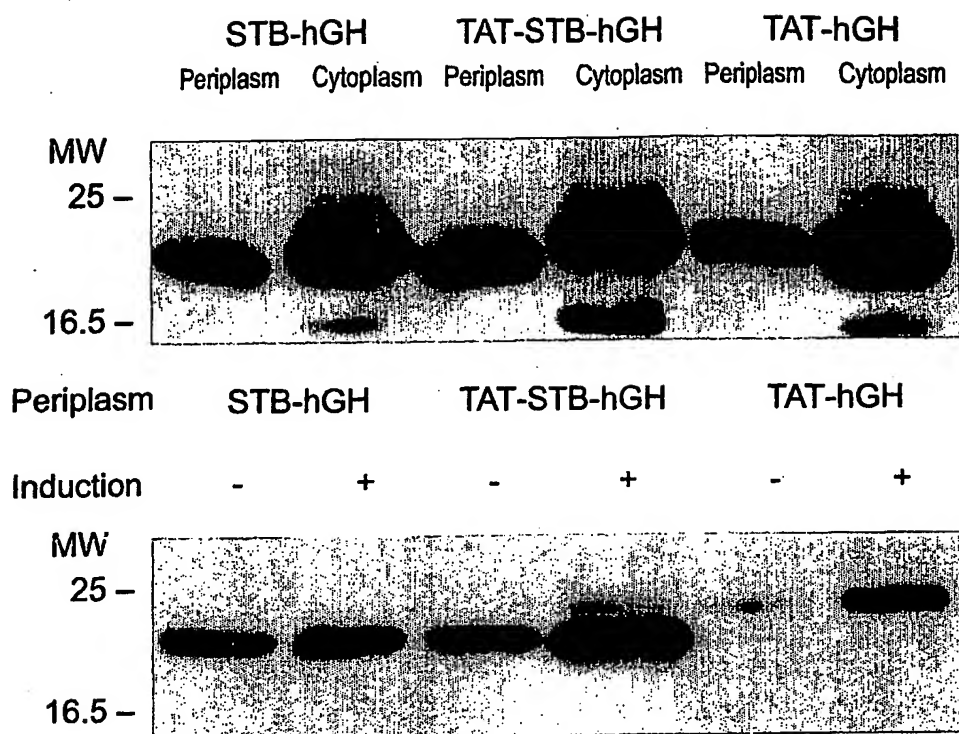


Fig. 4

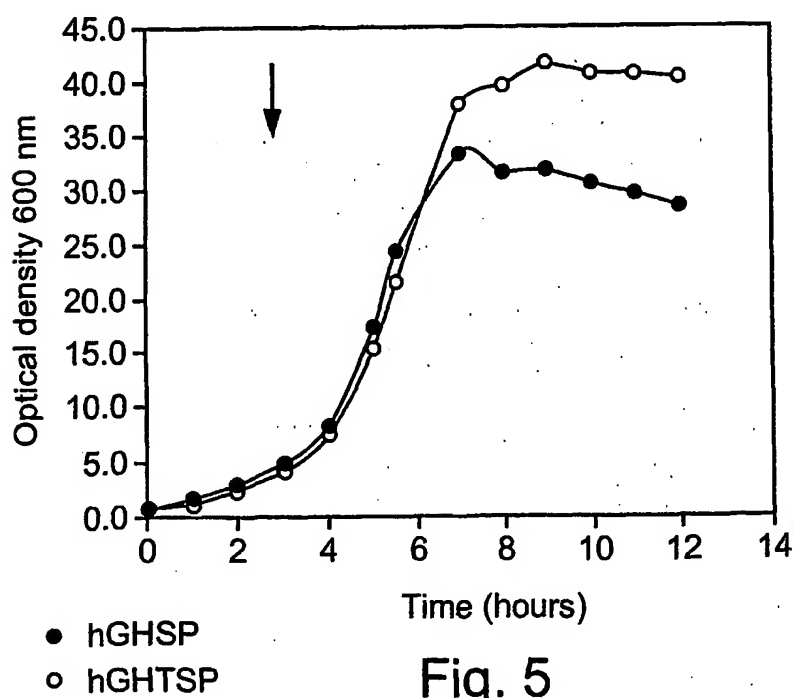


Fig. 5

I  
SEQUENCE LISTING

<110> Peleg, Yoav  
Pancer, Zeev

<120> PROKARYOTIC EXPRESSION CONSTRUCTS, METHODS OF GENERATING SAME AND  
METHODS OF USING SAME FOR EXPRESSION OF RECOMBINANT PROTEINS IN PROKARYOTIC  
EXPRESSION SYSTEMS

<130> 02/23924

<160> 18

<170> PatentIn version 3.1

<210> 1

<211> 576

<212> DNA

<213> Homo sapiens

<400> 1  
 ttccaacca ttccgttatc caggcttttt gacaacgcta tgctccgcgc ccatcgtctg 60  
 caccagctgg cctttgacac ctaccaggag ttgaagaag cctatatccc aaaggaacag 120  
 aagtattcat tcctgcagaa cccccagacc tccctctggt tctcagagtc tattccgaca 180  
 ccctccaaca gggaggaaac acaacagaaa tccaacctag agctgctccg catctccctg 240  
 ctgctcatcc agtcgtggct ggagcccgtg cagttcctca ggagtgtctt cgccaacagc 300  
 ctggtgtacg gcgcctctga cagcaacgct tatgacctcc taaaggacct agaggaaggc 360  
 atccaaacgc tgatggggag gctggaagat ggcagcccc ggactgggca gatcttcaag 420  
 cagacctaca gcaagttcga cacaactca cacaacgatg acgcactact caagaactac 480  
 gggctgctct actgcttcag gaaggacatg gacaaggctg agacattcct gcgcatcgtg 540  
 cagtgccgct ctgtggaggg cagctgtggc ttctag 576

2

&lt;210&gt; 2

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 2

gggctatgca ttcccaacca ttccgttatc caggc

35

&lt;210&gt; 3

&lt;211&gt; 36

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 3

acccgatcc ctagaagcca cagctgcct ccacag

36

&lt;210&gt; 4

&lt;211&gt; 72

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;



3

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 4

tatgaaaaag aatatcgcat ttcttcttgc atctatgttc gttttttcta ttgctacaaa 60

tgcctatgca tg 72

&lt;210&gt; 5

&lt;211&gt; 74

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 5

gatccatgca taggcatttg tagcaataga aaaaacgaac atagatgcaa gaagaaatgc 60

gatattcttt ttca 74

&lt;210&gt; 6

&lt;211&gt; 214

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; pSTB-hGH expression construct, open reading frame

&lt;400&gt; 6

Met	Lys	Lys	Asn	Ile	Ala	Phe	Leu	Leu	Ala	Ser	Met	Phe	Val	Phe	Ser
1			5					10					15		

4

Ile Ala Thr Asn Ala Tyr Ala Phe Pro Thr Ile Pro Leu Ser Arg Leu  
20 25 30

Phe Asp Asn Ala Met Leu Arg Ala His Arg Leu His Gln Leu Ala Phe  
35 40 45

Asp Thr Tyr Gln Glu Phe Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys  
50 55 60

Tyr Ser Phe Leu Gln Asn Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser  
65 70 75 80

Ile Pro Thr Pro Ser Asn Arg Glu Gln Thr Gln Gln Lys Ser Asn Leu  
85 90 95

Glu Leu Leu Arg Ile Ser Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro  
100 105 110

Val Gln Phe Leu Arg Ser Val Phe Ala Asn Ser Leu Val Tyr Gly Ala  
115 120 125

Ser Asp Ser Asn Val Tyr Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile  
130 135 140

Gln Thr Leu Met Gly Arg Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln  
145 150 155 160

Ile Phe Lys Gln Thr Tyr Ser Lys Phe Asp Thr Asn Ser His Asn Asp  
165 170 175

Asp Ala Leu Leu Lys Asn Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp  
180 185 190

Met Asp Lys Val Glu Thr Phe Leu Arg Ile Val Gln Cys Arg Ser Val  
195 200 205

5

Glu Gly Ser Cys Gly Phe  
210

&lt;210&gt; 7

&lt;211&gt; 51

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 7

catatgaaag gctatggccg caaaaaacgt cgccagcgtc gccgtggtgc a

51

&lt;210&gt; 8

&lt;211&gt; 47

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 8

ccacggcgac gctggcgacg ttttttgcg ccatagcctt tcatatg

47

&lt;210&gt; 9

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

<220>

<223> Single strand DNA oligonucleotide

<400> 9

aacatatgaa aggctatggc cgcaa

25

<210> 10

<211> 31

<212> DNA

<213> Artificial sequence

<220>

<223> Single strand DNA oligonucleotide

<400> 10

aaaggatcca ttagaagcca cagctgccct c

31

<210> 11

<211> 207

<212> PRT

<213> Artificial sequence

<220>

<223> pTAT-hGH expression vector, open reading frame

<400> 11

Met Lys Gly Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg Gly Ala  
1 5 10 15

7

Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu Arg  
                   20                  25                  30

Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe Glu  
                   35                  40                  45

Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn Pro  
                   50                  55                  60

Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn Arg  
                   65                  70                  75                  80

Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser Leu  
                   85                  90                  95

Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser Val  
                   100                  105                  110

Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr Asp  
                   115                  120                  125

Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg Leu  
                   130                  135                  140

Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr Ser  
                   145                  150                  155                  160

Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn Tyr  
                   165                  170                  175

Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr Phe  
                   180                  185                  190

Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe  
                   195                  200                  205

8

&lt;210&gt; 12

&lt;211&gt; 51

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 12

catatgaaag gctatggccg caaaaaacgt cgccagcgtc gccgtggcgc a

51

&lt;210&gt; 13

&lt;211&gt; 47

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 13

ccacggcgac gctggcgacg ttttttgcgg ccatagcctt tcatatg

47

&lt;210&gt; 14

&lt;211&gt; 51

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

9

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 14

tttcttcttg catctatggtt cgttttttct attgctacaa atgcctatgc a 51

&lt;210&gt; 15

&lt;211&gt; 51

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; Single strand DNA oligonucleotide

&lt;400&gt; 15

taggcatttg tagcaataga aaaaacgaac atagatgcaa gaagaaatgc g 51

&lt;210&gt; 16

&lt;211&gt; 224

&lt;212&gt; PRT

&lt;213&gt; Artificial sequence

&lt;220&gt;

&lt;223&gt; pTAT-STB-hGH epression vector, open reading frame

&lt;400&gt; 16

Met	Lys	Gly	Tyr	Gly	Arg	Lys	Lys	Arg	Arg	Gln	Arg	Arg	Arg	Gly	Ala
1				5				10						15	

Phe	Leu	Leu	Ala	Ser	Met	Phe	Val	Phe	Ser	Ile	Ala	Thr	Asn	Ala	Tyr
				20				25						30	

10

Ala Phe Pro Thr Ile Pro Leu Ser Arg Leu Phe Asp Asn Ala Met Leu  
35 40 45

Arg Ala His Arg Leu His Gln Leu Ala Phe Asp Thr Tyr Gln Glu Phe  
50 55 60

Glu Glu Ala Tyr Ile Pro Lys Glu Gln Lys Tyr Ser Phe Leu Gln Asn  
65 70 75 80

Pro Gln Thr Ser Leu Cys Phe Ser Glu Ser Ile Pro Thr Pro Ser Asn  
85 90 95

Arg Glu Glu Thr Gln Gln Lys Ser Asn Leu Glu Leu Leu Arg Ile Ser  
100 105 110

Leu Leu Leu Ile Gln Ser Trp Leu Glu Pro Val Gln Phe Leu Arg Ser  
115 120 125

Val Phe Ala Asn Ser Leu Val Tyr Gly Ala Ser Asp Ser Asn Val Tyr  
130 135 140

Asp Leu Leu Lys Asp Leu Glu Glu Gly Ile Gln Thr Leu Met Gly Arg  
145 150 155 160

Leu Glu Asp Gly Ser Pro Arg Thr Gly Gln Ile Phe Lys Gln Thr Tyr  
165 170 175

Ser Lys Phe Asp Thr Asn Ser His Asn Asp Asp Ala Leu Leu Lys Asn  
180 185 190

Tyr Gly Leu Leu Tyr Cys Phe Arg Lys Asp Met Asp Lys Val Glu Thr  
195 200 205

Phe Leu Arg Ile Val Gln Cys Arg Ser Val Glu Gly Ser Cys Gly Phe  
210 215 220



&lt;210&gt; 17

&lt;211&gt; 984

&lt;212&gt; DNA

&lt;213&gt; Artificial sequence

&lt;220&gt;

<223> An expression cassette containing: lambda phage gtl1 thermo-labile repressor under the control of a synthetic constitutive promoter; Phage lambda pL promoter; A multiple cloning site (MCS); and a transcriptional termination site (TT).

&lt;400&gt; 17

ggtacctcag ccaaactct cttcaggcca ctgactagcg ataactttcc ccacaacgga	60
acaactctca ttgcatggga tcattgggta ctgtgggttt agtgggttga aaaacacctg	120
accgctatcc ctgatcagtt tcttgaaggt aaactcatca cccccaagtc tggctatgca	180
gaaatcacct ggctcaacag cctgctcagg gtcaacgaga attaacattc cgtcaggaaa	240
gcttggett gaggcctgtg gtgcggtcac ggaattacct tcaacctcaa gccagaatgc	300
agaatcactg gcttttttgg ttgtgcttac ccattctctc gcattcacctt tggtaaaggt	360
tctaagctca ggtgagaaca tccctgcctg aacatgagaa aaaacagggg actcatactc	420
acttctaagt gacggctgca tactaaccgc ttcatacatc tcgtagattt ctctggcgat	480
tgaagggcta aattcttcaa cgctaacttt gagaattttt gtaagcaatg cggcggtata	540
agcatttaat gcattgatgc cattaaataa agcaccaacg cctgactgcc ccatcccat	600
cttgctctcg acagattcct gggataagcc aagttcattt ttcttttttt cataaattgc	660
tttaaggcga cgtgcgtcct caagctgctc ttgtgttaat ggtttctttt ttgtgctcat	720
ctcgagcctc ctatagttag tcgtattata ctatgccgat gttattgtca aaagcttgat	780
atcgaattct gcaaaaaata aattcatata aaaaacatac agataaccat ctgcggtgat	840
aaattatctc tggcgggtgtt gacataaata ccactggcgg tgatactgag cacatcagca	900

12  
ggactctaga gcggccgcgg gcccgtcgac atcgatctgc agcccgggat ccactagtgg 960  
cccacccgaa ggtgagccga gctc 984

<210> 18

<211> 11

<212> PRT

<213> Artificial sequence

<220>

<223> TAT-derived peptide

<400> 18

Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg  
1 5 10

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
16 January 2003 (16.01.2003)

PCT

(10) International Publication Number  
**WO 2003/004599 A3**

(51) International Patent Classification<sup>7</sup>: C12P 21/04,  
C12N 15/00, 15/09, 15/63, 15/70, 15/74, C07H 21/02,  
21/04

(21) International Application Number:  
PCT/IL2002/000540

(22) International Filing Date: 2 July 2002 (02.07.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/301,805 2 July 2001 (02.07.2001) US

(71) Applicant (for all designated States except US): IN-  
SIGHT BIOPHARMACEUTICALS LTD. [IL/IL];  
Rabin Science Park, P.O. Box 2128, 76 121 Rehovot (IL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): PELEG, Yoav  
[IL/IL]; 18 Neve Alon Street, 76 495 Rehovot (IL).  
PANCER, Zeev [IL/IL]; 3/22 Rav Ashi Street, 69 395 Tel  
Aviv (IL).

(74) Agent: G. E. EHRLICH (1995) LTD.; 28 Bezalel Street,  
52 521 Ramat Gan (IL).

(81) Designated States (*national*): AE, AG, AL, AM, AT (util-  
ity model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA,  
CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (util-  
ity model), DE, DK (utility model), DK, DM, DZ, EC, EE  
(utility model), EE, ES, FI (utility model), FI, GB, GD, GE,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ,  
LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN,  
MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD,  
SE, SG, SI, SK (utility model), SK, SL, TJ, TM, TN, TR,  
TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK,  
TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,  
GW, ML, MR, NE, SN, TD, TG).

Published:

— with international search report

(88) Date of publication of the international search report:  
18 March 2004

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

WO 2003/004599 A3

(54) Title: METHODS OF MAKING AND USING EXPRESSION CONSTRUCTS FOR THE EXPRESSION OF RECOMBI-  
NANT PROTEINS

(57) Abstract: A method of producing a fusion polypeptide in, and purifying the fusion polypeptide from, bacteria, the method  
comprises introducing into the bacteria an expression construct encoding the fusion polypeptide which comprises a TAT-derived  
peptide, and a protein-of-interest, whereby the TAT-derived peptide serves for transport of the fusion polypeptide from the bacterial  
cytoplasm to the periplasm. The fusion polypeptide is then substantially exclusively purified from the bacterial periplasm.

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/IL02/00540

**A. CLASSIFICATION OF SUBJECT MATTER**

IPC(7) : C12P 21/04; C12N 15/00, 15/09, 15/63, 15/70, 15/74; C07H 21/02, 21/04

US CL : 435/69.7, 71.1, 320.1, 471; 536/23.1

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.7, 71.1, 320.1, 471; 536/23.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
Please See Continuation Sheet**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	US 6,221,355 B1 (DOWDY) 24 April 2001 (24.04.2001), see entire document.	1-108
A, P	US 6,485,977 B1 (COLLMER et al.) 26 November 2002 (26.11.2002), see entire document.	1-108
A, P	US 2003/0018001 A1 (HEANEY et al.) 23 January 2003 (23.01.2003), see entire document.	1-108
A, P	US 2003/00502243 A1 (TYMIANSKI) 13 March 2003 (13.03.2003), see entire document.	1-108

☐ Further documents are listed in the continuation of Box C.☐ See patent family annex.

Special categories of cited documents:	
* "A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

07 April 2003 (07.04.2003)

Date of mailing of the international search report

30 JUN 2003

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

Konstantina Katcheves

Telephone No. (703) 308-1235

## INTERNATIONAL SEARCH REPORT

PCT/IL02/00540

**Continuation of Item 4 of the first sheet:**

**Examiner:** The current title is longer than seventeen words. Thus, the title fails to satisfy PCT Rule 4.3.

**New Title:** Methods of making and using expression constructs for the expression of recombinant proteins.

**Continuation of B. FIELDS SEARCHED Item 3:**

**EAST, STN, CAPLUS, BIOSIS, MEDLINE**

**search terms:** expression construct, vector, TAT peptide, TAT-derived peptide, YGRKKRRQRRR

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record.**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**